

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 14:06:06 ; Search time 17.59 Seconds  
(without alignments)  
1001.557 Million cell updates/sec

Title: US-09-464-099A-70  
Perfect score: 2288  
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

al number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282	99.7	455	1 AROA_AGRSP	Q9r4e4 agrobacteri
2	1900.5	83.1	449	1 AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1 AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1 AROA_SVNY3	Q59975 synecocyst
5	806.5	35.2	430	1 AROA_LACLA	Q9ceu0 lactococcus
6	806	35.2	443	1 AROA_BACNO	Q46550 bacteroides
7	803.5	35.1	427	1 AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1 AROA_LACLC	P43905 lactococcus
9	785	34.3	428	1 AROA_BACSU	P20691 bacillus su
10	766.5	33.5	431	1 AROA_AQUAE	Q67494 aquifex aeo
11	695	30.4	410	1 AROA_THEMEA	Q9wy10 thermotoga
12	590	25.8	428	1 AROA_CWJJE	P52312 campylobact
13	577	25.2	430	1 AROA_STAAU	Q05615 staphylococ
14	521.5	22.8	429	1 AROA_HELPJ	Q9zkf7 helicobacte
15	518.5	22.7	429	1 AROA_HELPY	P56197 helicobacte
16	496	21.7	419	1 AROA_METTH	Q26860 methanobact
17	459.5	20.1	429	1 AROA_METJA	Q57925 methanococc
18	408	17.8	439	1 AROA_DEIRA	Q9rvd3 deinococcus
19	403	17.6	416	1 AROA_ARCFU	O28775 archaeoglob
20	402.5	17.6	427	1 AROA_AERPE	Q9yek9 aeropyrum p
21	402.5	17.6	427	1 AROA_YEREN	P19688 versinia en
22	395.5	17.3	428	1 AROA_YERPE	O60112 versinia pe
23	387.5	16.9	432	1 AROA_HAESO	P52310 haemophilus
24	378	16.5	427	1 AROA_ECOLI	P07638 escherichia
25	376	16.4	427	1 AROA_SHISO	Q9zif7 shigella so
26	373.5	16.3	410	1 AROA_PYRAB	Q9v1h1 pyrococcus
27	373	16.3	432	1 AROA_PASHA	P54220 pasteurella
28	372	16.3	427	1 AROA_KLEPN	P24497 klebsiella
29	372	16.3	440	1 AROA_PASMU	Q04570 pasteurella
30	371	16.2	427	1 AROA_SHIDY	O87006 shigella dy
31	365.5	16.0	463	1 AROA_BURPS	P39915 burkholderi
32	365	16.0	428	1 AROA_EDWIC	Q9x4h2 edwardsiell
33	363.5	15.9	432	1 AROA_HAEIN	Q03421 haemophilus

34	357	15.6	427	1 AROA_SALTY	P07637 salmonella
35	348	15.2	427	1 AROA_SALGL	P22299 salmonella
36	347	15.2	427	1 AROA_SALTI	P19786 salmonella
37	323.5	14.1	516	1 AROA_PETHY	P11043 petunia hyb
38	321.5	14.1	1588	1 ARO1_YEAST	P08566 s pentafunc
39	317	13.9	441	1 AROA_CHLMU	Q9pk28 chlamydia m
40	315.5	13.8	518	1 ARO1_TOBAC	P23981 nicotiana t
41	314.5	13.7	440	1 AROA_CHLTR	O84371 chlamydia t
42	313	13.7	428	1 AROA_BUCAP	Q59178 buchnera ap
43	310.5	13.6	520	1 AROA_LYCES	P10748 lycopersico
44	308.5	13.5	1573	1 ARO1_SCHPO	Q9p7r0 s pentafunc
45	306.5	13.4	445	1 AROA_CHLPN	Q9z6m0 chlamydia p

ALIGNMENTS

RESULT	1
AROA_AGRSP	
ID AROA_AGRSP	STANDARD; PRT; 455 AA.
AC Q9R4E4;	
DT 30-MAY-2000 (Rel. 39, Created)	
DT 30-MAY-2000 (Rel. 39, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).	
DE enolpyruvylshikimate-3-phosphate synthase	
GN AROA.	
OS Agrobacterium sp. (strain CP4).	
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;	
OC Rhizobiaceae; Agrobacterium.	
OX NCBI_TaxID=361;	
RN [1]	
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.	
RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;	
RT "Glyosphate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";	
RL Patent number US5633435, 27-MAY-1997.	
RN [2]	
RP SEQUENCE OF 1-15.	
RX MEDLINE=96182485; PubMed=8598558;	
RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,	
RA Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,	
RA Fuchs R.L., Padgett S.R.;	
RT "The expressed protein in glyphosate-tolerant soybean, 5-enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp. strain CP4, is rapidly digested in vitro and is not toxic to acutely gavaged mice.";	
RT J. Nutr. 126:728-740(1996).	
RL -I- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate - phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.	
CC -I- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).	
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).	
CC -I- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in glyphosate-tolerant soybean, canola, cotton and maize by Monsanto. Developed to provide new weed-control options for farmers.	
CC Expression of this protein in plants imparts high levels of glyphosate tolerance.	
CC -I- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.	
DR InterPro: IPR001986; EPSP_synthase.	
DR Pfam: PF00275; EPSP_synthase; 1.	
DR PRODOM: PD001867; EPSP_synthase; 1.	
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.	
DR PROSITE: PS00885; EPSP_SYNTHASE_2; 1.	
KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance; Genetically modified food.	
FT CONFLICT 2 2 S -> L (IN REF. 2).	
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;	

Query Match 99.7%; Score 2282; DB 1; Length 455;  
Best Local Similarity 99.8%; Pred. No. 1.le-134;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MLHGASSRATARKSSGLSGT	VRIPCDKSI	SHRSEFMFG	LAGSETRIT	GLLEGEDV	INTG	60
DB	1	MSHGASSRATARKSSGLSGT	VRIPCDKSI	SHRSEFMFG	LAGSETRIT	GLLEGEDV	INTG	60
QY	61	KAMQANGARIRKEGDTWII	DGVNGGLLA	PEAPLDFG	NAATGCR	LTMTGLV	GVYDFDSTFI	120
DB	61	KAMQANGARIRKEGDTWII	DGVNGGLLA	PEAPLDFG	NAATGCR	LTMTGLV	GVYDFDSTFI	120
QY	121	GDASLTKRPMGRVNLREL	MGVQVKS	EDGDR	LPVTLR	GP	KTPPTI	TYRVPMA
DB	121	GDASLTKRPMGRVNLREL	MGVQVKS	EDGDR	LPVTLR	GP	KTPPTI	TYRVPMA
QY	181	LLAGLNTPGITTVIEP	IMTRDHT	EKMLQ	FGANL	TVETD	ADGVRT	IRLEGR
DB	181	LLAGLNTPGITTVIEP	IMTRDHT	EKMLQ	FGANL	TVETD	ADGVRT	IRLEGR
QY	241	VPGPSSTAPPLVAALL	VP	GS	DVTILN	VL	MNPT	RTGL
DB	241	VPGPSSTAPPLVAALL	VP	GS	DVTILN	VL	MNPT	RTGL
QY	301	VADLRVRSSTLKGVT	YEDRAP	SMID	EYPT	IL	VA	AAFA
DB	301	VADLRVRSSTLKGVT	YEDRAP	SMID	EYPT	IL	VA	AAFA
QY	361	VANGLKLVGDCDE	GETSL	VYVR	GR	PD	KG	LGN
DB	361	VANGLKLVGDCDE	GETSL	VYVR	GR	PD	KG	LGN
QY	421	VTYDDATMIATSP	PEFMD	L	MAG	L	GA	K
DB	421	VTYDDATMIATSP	PEFMD	L	MAG	L	GA	K
RESULT	2							
ARO	ARO	PSES2	STANDARD;	PRT;	449	AA.		
AC	P56952;							
DT	30-MAY-2000	(Rel. 39, Created)						
DT	30-MAY-2000	(Rel. 39, Last sequence update)						
DT	16-OCT-2001	(Rel. 40, Last annotation update)						
DE	3-phosphoshikimate	1-carboxyvinyltransferase (EC 2.5.1.19) (5-						
DE	enolpyruvylshikimate	3-phosphate synthase) (EPSP synthase) (EPSPS).						
GN	ARO							
OS	Pseudomonas sp. (strain PG2982), and							
OC	Achromobacter sp. (strain LBAA).							
OC	Bacteria; Proteobacteria.							
OX	NCBI_TaxID-308, 129026;							
RN	[1]							
SP	SEQUENCE FROM N.A., AND SEQUENCE OF 2-16.							
SP	Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;							
SP	"Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";							
RL	Patent number US5633435, 27-MAY-1997.							
CC	-1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =							
CC	phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.							
CC	-1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHLORISIMATE WITHIN							
CC	THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).							
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).							
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).							
CC	-1- MISCELLANEOUS: RESISTANT TO THE ANTIBIOTIC GLYPHOSATE.							
CC	-1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.							
DR	InterPro: IPR001986; EPSP_synthase.							
DR	Pfam: PF00275; EPSP_synthase; 1.							
DR	Prodom; P0001667; EPSP_synthase; 1.							
DR	PROSITE; PS00104; EPSP_SYNTHASE_1; 1.							
DR	PROSITE; PS00885; EPSP_SYNTHASE_2; 1.							
KW	Aromatic amino acid biosynthesis; Transferase; Herbicide resistance.							
SO	SEQUENCE 449 AA; 47297 MW; 447F213ECCAEFC1 CRC64;							

Query Match	83.1%	Score 1900.5;	DB 1;	Length 449;
Best Local Similarity	82.9%	Pred. No. 4.8e-111;		
Match 36	373;	Conservative	32;	Mismatches 42;
				Indels 3;
				Gaps 1;

QY	1	MLHGASSRPATARKSSGLSETVRIPGDKSIHSRSPFMGGLASGETRITGLLGEGDVINTG	60
Db	1	MSHSAPKPTARRSALTSALGTEIRIPGDKSIHSRSPFMGGLASGETRITGLLGEGDVINTG	60
QY	61	KAMQAMGARIRKEGDTWIIDGVNGVNGLLAPELLDFGNAAATCRLTMGLVGVYDFDSTFI	120
Db	61	RAMQAMGAKIRKEGDWIIINGVNGVNGLLQPEAALDFGNAGTCARLTMTGLVGYDMKTSFI	120
QY	121	GDASLTKRPMGRVNLPRMGVQVKSESDGRPLPVTLRGPKTPTPTIYRVPMAAQVKSVA	180
Db	121	GDASLSKRPMGRVNLPRMGVQVEAAGDGRPLTTLIGPKTANPTIYRVPMAAQVKSVA	180
QY	181	LLAGLNTPGTITTVIEPIMTRDHTKMLQGFALNTVETDADGVRTIRLEGRGKLTGQVTD	240
Db	181	LLAGLNTPGTITTVIEPIMTRDHTKMLQGFADLTVETDKGVHRHIRTGQGLVGQTD	240
QY	241	VPGDSPSTAPPLVAALLVPGSDVTIILNLMNPRRTGLILTLQEMGADIEVINPRLAGGED	300
Db	241	VPGDSPSTAPPLVAALLVPGSDVTIRNLMNPRRTGLILTLQEMGADIEVINPRLAGGED	300
QY	301	VADLRVRSSTLKVTVPEDRAPSMIDEPYPLAVAAAFACGATVMNGLEELRVKESDRLSA	360
Db	301	VADLRVRASLKVGVVPPERAPSMIDEPYPLVLAATAFAEGETVMDGLDELVRKESDRLAS	360
QY	361	VANGLKTLNGVDCDEGETSLVVVRGPRDGKGLGNASGAAVATHLDRHTAMSLVMGLVSENP	420
Db	361	VARGLEANGVDCTEGEMSLAVRGRPDGKGLG---GGTVATHLDRHTAMSLVMGLAAEKP	417
QY	421	VTYDDATMIATSPERPMDLNLGAKLIELS	450
Db	418	VTYDDSNMIATSPERPMDMPLGAKLIELS	447
RESULT 3			
AD	ARO_A_BACHD	STANDARD;	PRT; 431 AA.
AC	Q9KCA6;		
DT	01-MAR-2002 (Rel. 41, Created)		
DT	01-MAR-2002 (Rel. 41, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-phosphoshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).		
GN	ARO_A OR AOE OR BH1667.		
OS	Bacillus halodurans.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCBI_TaxID=86665;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=C-125 / JCM 9153;		
RA	MEDLINE=20512582; PubMed=11058132;		
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,		
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,		
RA	Horikoshi K.;		
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus		
RT	halodurans and genomic sequence comparison with Bacillus subtilis.";		
RL	Nucleic Acids Res 28:4317-4331(2000).		
CC	-!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =		
CC	phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.		
CC	-!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE		
CC	BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).		
CC	-!- SUBUNIT: Monomer (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).		
CC	-!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-		
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outsta-		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in the		
CC	modified and this statement is not removed. Usage by and for comm-		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/ann">http://www.isb-sib.ch/ann</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		

DR EMBL; AP001512; BAB05386.1; -  
DR InterPro; IPR001986; EPSP\_synthase.  
DR Pfam; PF00275; EPSP\_synthase; 1.  
DR ProDom; PD001867; EPSP\_synthase; 1.  
DR PROSITE; PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE; PS00885; EPSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
SQ SEQUENCE 431 AA; 45485 MW; 12F4PBE7BA0743D CRC64;  
  
Query Match 38.6%; Score 884; DB 1; Length 431;  
Best Local Similarity 44.2%; Pred. No. 5e-48;  
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;  
  
QY 15 SSGLSGTVRIPGDKSISRSEFMFGGLASGETRITGLLEGEDVINTGKAMQANGARIRK 74  
DB 10 AKGLGKTVIRPEDKKSISRHVMFGALAGTITVEGFLPGADCLSTISCFKLGVSIEQAE 69  
QY 75 DTWIIDGVNGGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFGDASLTKRPMGRVL 134  
DB 70 ERVTVKGGWGLREPSDILDVNSGTTTLLILGILSLPFSHVIIGDSIGKRPMKRV 129  
QY 135 NPLRMGVQVKSSED-GRDLPVTLRGPKTPPTTYRVPMSAQVKSAVLLAGLNTPGITTV 193  
DB 130 EPLKSMGAQIDGRDHCNLTPLSIRGGQL-KGIDFSPVSAQMKSAILLAGLRAEGKTSV 188  
QY 194 IEPINTRDHTEKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSSTAFPLV 253  
DB 189 TEPAKTRDHTERTMLEAFGVNI-----EKDGL-TVSIEGGQMLTGQHVVPVPGDISSAAFFLV 243  
QY 254 AALLVPGSDVITLNLVMPNPTRTGLTLTQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313  
DB 244 AGAMVPHSRITLVNGINPTFRAGILEVUKMGATUAMENRVOGGEPAVDLTIETSVLQG 303  
QY 314 VTPEDRAPSMIDEYPIILAVAAAFAGATVNMGLEELRVKESDRLSAVANGKLKGVDCD 373  
DB 304 VEIGGDIIPRLIDEPIITAVLATQASGRVTIKDAELVKETNRDITVVSSELTKLGAISH 363  
QY 374 EGETSLVVRGRPDGKGLNAGSAVAATHLDRHANSFLVGLVNSPVTVDATMIATSF 433  
DB 364 ATDDGMIETGTPPLKG-----GVTSSHGDRHGMAMAIALLAELKPVTVTEGTEAIAVSY 418  
QY 434 PEFMD 438  
DB 419 PSFED 423

RESULT 4

AC AROA\_SYN3 STANDARD; PRT; 447 AA.  
QY Q59975; Q59974;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
GN AROA OR SFR044.  
OS Synchocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94299161; PubMed=8026753;  
RA dalla Chiesa M., Mayes S.R., Maskell D.S., Nixon P.J., Barber J.;  
RT "An aroA homologue from Synchocystis sp. PCC 6803.";  
RL Gene 144:145-146(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugliura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraiki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
RN [3]  
RP SEQUENCE OF 103-194 FROM N.A.  
RX MEDLINE=93307506; PubMed=7686511;  
RA Mayes S.R., dalla Chiesa M., Zhang Z., Barber J.;  
RT "The genes aroA and trnQ are located upstream of psbO in the  
RT chromosome of Synchocystis 6803.";  
RL FEBS Lett. 325:255-261(1993).  
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISHIKIMATE WITHIN  
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
CC EMBL; X75325; CAA53074.1; -  
CC EMBL; D90914; BAA18477.1; -  
CC EMBL; X72784; CAA51291.1; -  
CC InterPro; IPR001986; EPSP\_synthase.  
CC Pfam; PF00275; EPSP\_synthase; 1.  
CC ProDom; PD001867; EPSP\_synthase; 1.  
CC PROSITE; PS00104; EPSP\_SYNTHASE\_1; 1.  
CC PROSITE; PS00885; EPSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
SQ SEQUENCE 447 AA; 47046 MW; 2F3C8AD26B5A7BCE CRC64;  
  
Query Match 37.9%; Score 867.5; DB 1; Length 447;  
Best Local Similarity 48.1%; Pred. No. 5.5e-47;  
Matches 203; Conservative 62; Mismatches 144; Indels 13; Gaps 7;  
  
QY 18 LSGTVRIPGDKSISRSEFMFGGLASGETRITGLLEGEDVINTGKAMQANGARIRK-EGDT 76  
DB 26 LTGRLRVPGDKSISRHVMFGALAGTITVEGFLPGADCLSTISCFKLGVSIEQAE 85  
QY 77 WIIDGVNGGGLLAPEAPLDFGNAATGRLTMGLV-CVYDFDSTFGDASLTKRPMGRVLN 135  
DB 86 IIVQGRGLGQLQEPSTVLDAGNSGTTMRLMGLLAGQKCLFTVTGDDSLRHPSRVIQ 145  
QY 136 PLREMGVQV-KSEDDRLPVTLRGPKTPPTTYRVPMSAQVKSAVLLAGLNTPGITTVI 194  
DB 146 PLQMGAKIWARSKNGKFAFLAVQGSQ-L-KPIHYHSPIASAQVKSCLLAGLITGEGTTV 204  
QY 195 EPIPTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSSTAFPLVA 254  
DB 205 EPALSRDHSERMLQAFGAKLTID---PVTHSVTVHGAHLTGQVVVPGDIISSAAFVLA 261  
QY 255 ALLVPGSDVITLNLVMPNPTRTGLTLTQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 314  
DB 262 ASILPSELVENVGINTPTRTGVLVLAQMGADITPENERLVTGEPVADLRVRAHSHQGC 321  
QY 315 TYPEDRAPSMIDEYPIILAVAAAFAGATVNMGLEELRVKESDRLSAVANGKLKGVDCD 374  
DB 322 TFGGEIIPRLIDEPIITAVLATQASGRVTIKDAELVKETNRDITVVSSELTKLGAISH 381  
QY 375 GETSLVVRGRPDGKGLNAGSAVAATHLDRHANSFLVGLVNSPVTVDATMIATSP 434  
DB 382 FDDGLEIQG---GSPL---OQAEVDSLTDRHAMAIALAALGSGGQTIINRAAAAISSP 435  
QY 435 EF 436

DB 436 EF 437

## RESULT

```

AROA_LACIA          STANDARD;          PRT;      430 AA.
ID
AC Q9CEU0;
AC 16-OCT-2001 (Rel. 40, Created)
AC 16-OCT-2001 (Rel. 40, Last sequence update)
AC 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
DE AROA OR LL1744
GN Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
OX [1]
RP SEQUENCE FROM N.A.
RP STRAIN=IL1403;
MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winker P., Mauger S., Jaillon O., Malarme K.,
RA Wellesenbach J., Ehrlich S.D., Sorokin A.;
RA "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL: AE006404; AAK05842.1;
CC InterPro: IPR001986; EPSP_synthase.
CC Pfam: PF00275; EPSP_synthase; 1.
CC ProDom: PD001867; EPSP_synthase; 1.
CC PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
CC PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferase; Complete proteome.
CC SOURCE: 430 AA: 45958 MW: CB216F07AA4EE799 CRC64;
CC -----

```

[illegible]

## RESULT 6

```

AROA_BACNO
ID AROA_BACNO STANDARD; PRT; 443 AA.
AC 046550;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) [5-
DE enolpyruvylshikimate-3-phosphate synthase] (EPSP synthase) (EPSPS).
GN AROA.
OS Bacteroides nodosus [Dichelobacter nodosus].
OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OC Dichelobacter.
OC NCBI_TaxID=870;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=VCS1001.
RC MEDLINE=94320795; PubMed=8045432;
RA Alm R.A., Dalrymple B.P., Mattick J.S.;
RT "Sequencing and expression of the aroA gene from Dichelobacter
RT nodosus.";
RL Gene 145;97-101(1994).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29339; CAA82544.1; .
CC InterPro; IPR001986; EPSP_synthase.
CC Pfam; PF00275; EPSP_synthase; 1.
CC ProDom; PD001867; EPSP_synthase; 1.
CC PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
CC PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferase.
CC SEQUENCE 443 AA; 47444 MW; BZ243277ADEFB95 CRC64;
SQ

```

Query Match 35.2%; Score 806; DB 1; Length 443;  
Best Local Similarity 42.2%; Pred. No. 3.6e-43;  
Matches 185; Conservative 74; Mismatches 159; Indels 10; Gaps 3;



```
QY 76 TWIIDGVNGGLLAPDLFGNAATGCBRTMCLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db 72 IVTIRGVGFLGQPPKAPLNMONGSTSMRLLAGILAAQRFESVLCODESLKRPQMORIIT 131
QY 136 PLEMGVOVKSEGDRLPVTLRGPKPTPTTYRVPMAAQVKSVAVLGAGLNPFGITVIE 195
Db 132 PLVOMGAKIVSHSNFTAPLHIS-RLPTGIDVALPLPSAQLKSLILAGLADGTRLRHT 190
QY 196 PIWDRHTEKMLGGGANITVETDADGVRTIRLEGRKKTIGQVIDVPGDPSSSTAPFLVAA 255
Db 191 CGISRDHTEKMLPFGGALIEKKE-----QIIIVTGQKLGHCGLDITVGDLSAAAFPMVAA 245
QY 256 LVPVSGDVTILVNMNPTGTLITLQEMGADIEVNINPLAGGEDVADLRVRSSTLKVTT 315
Db 246 LIAPRAEVIRNVGINPTRAAITLLOKMGGRTELHHRQFWGAEPVADIVVHHSKURGIT 305
QY 316 VPEDRAPSMIDEPYILAVAAFAEGATVNMGLFELRVKESDRLSAVANGKLGNGVDCDEG 375
Db 306 VAPEWIANAIDELPIFFIAACAEGTTFVGNLSLSELRVKESDRLAAMAQNLQTLGVACDVG 365
QY 376 ETSLVVGRPDCKGLGNASGAAYATHLDHRIAMSFYMGVLVSENPTVDDATMIATSFPE 435
Db 366 ADFIHYGRSDHQFL-----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMFQ 421
QY 436 FMDLMAGLGAKEIELSDTK 453
Db 422 FRDFAAAGMNVGERDAK 439

RESULT 7
AROA_STRPN STANDARD; PRN; 427 AA.
AC O9SA400:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR SPI371.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX STRAIN=NCIMB 40794 / 0100993;
RX MEDLINE=20069365; PubMed=10601870;
Du W., Wallis N.G., Mazzulla M.J., Chalker A.F., Zhang L., Liu W.-S.,
Kallender H., Payne D.J.;
"Characterization of Streptococcus pneumoniae 5-enolpyruvylshikimate
3-phosphate synthase and its activation by univalent cations.";
RL Eur. J. Biochem. 267:222-227(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
  pneumoniae.";
RL Science 293:498-506(2001).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
  phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
  THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
```

```
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF169483; AAD45819.1; -.
DR TIGR; SP1371; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT CONFLICT 13 13 S -> I (IN REF. 1).
FT CONFLICT 71 71 I -> V (IN REF. 1).
FT CONFLICT 201 201 K -> Q (IN REF. 1).
FT CONFLICT 303 303 G -> C (IN REF. 1).
SQ SEQUENCE 427 AA; 45766 MW; 45CE6F4D0D1C7B70 CRC64;

Query Match 35.1%; Score 803.5; DB 1; Length 427;
Best Local Similarity 42.8%; Pred. No. 4.9e-43;
Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIPGDKSIHRSFEMFGGLASGETRITGLLEGEDVINTKAMQAMGARIRKSGDTW 77
Db 10 LHGSIRVPGDKSIHRSIIIFGSLAEGTKVYDLRGEDVLSTMQVFRDLGVEIEDKQVI 69
QY 78 IDGVNGGLLAPDLFGNAATGCRITMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
Db 70 TIQGVGMAGLKAPQNALNMGNSCTIRLISGLVLAGADFEMFGDDSLSKRPMDRVTLPL 129
QY 138 REMGVQVKS- DGRPLPVTLRGPKPTPTTYRVPMAAQVKSVAVLGAGLNPFGITVIEP 196
Db 130 KMGVSIQSQTERDLPRLKTKNLRPIHYELPIASQVKSALMFAALQAKGESVIEIK 189
QY 197 IFRDHTKMLQGFANLTVETDADGVRTIRLEGRKKTIGQVIDVPGDPSSSTAPFLVAA 256
Db 190 EYTRNTEKMLQFGHLSV----DG-KKITVGPKLTKGVVPGDISSAFLVAGL 244
QY 257 LVPVSGDVTILVNMNPTGTLITLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKV 314
Db 245 IAPNRLVLQNVGINETRTGIIDVIRAMGKLEITEIDPVAKS---ATLIVESDLKGT 300
QY 315 TYPEDRAPSMIDEPYILAVAAFAEGATVNMGLFELRVKESDRLSAVANGKLGNGVDCDE 374
Db 301 EIGGALIPRLIDELPIIALLATQAGVTVIKDAEELKVKETDRIQVYVADALNSMGADITP 360
QY 375 GETSLVVRGPRDCKGLGNASGAAYATHLDHRIAMSFYMG-VLSENPTVDDATMIATSF 433
Db 361 TADGMIKKGK-----SALHGARNVTFGHRIGCMTIAALLVADGEVELDRAEINTSY 414
QY 434 PEPMDLMAGL 443
Db 415 PSFDDLES 424

RESULT 8
AROA_LACLC STANDARD; PRN; 430 AA.
AC P43905;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
  enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
```

Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).  
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Lactococcus.  
NCBI\_TaxID=1359;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=MGI1363 / FJ5876;  
RX MEDLINE=95124293; PubMed=7823907;  
RC Griffin H.G., Gasson M.J.;  
RA "Genetic aspects of aromatic amino acid biosynthesis in Lactococcus  
RT lactis.";  
RL Mol. Gen. Genet. 246:119-127(1995).  
CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
CC -|- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
-----  
this SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to licensed@isb-sib.ch).  
-----  
EMBL: X78413; CAA53180.1; .  
DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_SYNTHASE\_1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; transferase.  
SQ SEQUENCE 430 AA; 45804 MW; C5A197A49072C9D5 CRC64;

Query Match 35.1%; Score 802.5; DB 1; Length 430;  
Best Local Similarity 41.2%; Pred. No. 5,7e+43;  
Matches 181; Conservative 79; Mismatches 158; Indels 21; Gaps 7;

QY 15 SSGLSGTVRIPCDKSIHSRSPFMGLASGETRITGLLEGEDVINTGKAMQAGRIKEG 74  
DB 7 SOGLKGRLKVPDGDKSIHSRSPFMGLASGETRITGLLEGEDVINTGKAMQAGRIEIDDG 66

QY 75 DTWIIDGVNGGLLAPEAPLDGNAATCGRTLTMGLVGVDFTDFIGDASLTKKPMGRVL 134  
DB 67 QVITVHGOGISKLKPEKALDMGNSTGTRLLSLGIALGPETALFLGDSDLSKRPMDRVA 126

QY 135 NPLREMGVQVKSE-DGDRLVPTLRPKPTPTTYRVPMVASAQKVSALLAGLAGNPTITVV 193  
DB 127 TPLQMGAIEIVGQTDKVKLPMTIKCSAHLKAIDYLTPVASQKVSAVIFAALQAEGTKV 186

QY 194 IEPIMTRDHTEKMLGFGANLTVEDTDAGVRTIRLEGRGKLTGTVIDPGDPSSAFPLV 253  
DB 187 VEKEXTRSHEMLVQFGEIITVD----KTPIYPGGOKLGQEVTPVGDLSAFAFLV 241

QY 254 RALLVPGSDVTILNLMNPTRGLTILTQEMGADIEVINPRLAGGED----VADLRVRSS 309  
DB 242 AGLVENSGLIILENVGINETRGTGLEIVQAMGQGLEILE-----QDEVAKAATLKVRAS 295

QY 310 TLKGVTVPEDRAPSMIDEYPIIAVAAAFAEGATVMNGLEELRKVSKSDRLSAVANGKLNG 369  
DB 296 QLKGTIEISGDLPRILDEIPLIALLTAQTAEKGTIIIDAELAEKVKETDRIAVVADALNSMG 355

QY 370 VDCDEGETSLVVRGPDGKGLNGASAAVATHLDHRIAMSFLVMGLVSN-PVTVDDATM 428  
DB 356 ANIEPTDGMCIQQ---GTKL-HAPENAINTLCDHRIGMMVAIALLLVENGEIELERAEA 411

QY 429 IATSFPEFMDLAGLAKI 447  
DB 412 IOTSYPSFFDLEKLSGNL 430

```
Db 239 AAGAMPNSRIVLKNVGLNPTRTGIIDVLONMGAKLEIKPSADSGAEPVGDLLIETSSLK 298
Qy 313 GVTVPEDRAPSMIDYPIILAVAAFAEAGATVNMGLEELRVKESDRLSAVANGKLNQVDC 372
Db 299 AVEIGGDIPIRLIDIPILALATQAEQTVTKDAELKVKTETNRDITVSELRLGABEI 358
Qy 373 DSGETSLVVRGPDCKGLGNAGAAVATHLDRHIAVMSFLVNGLVSENPVYDDATMIATS 432
Db 359 EPTADGMKYVQQTUKG-----GAAVSSHGDHRIGMMLGIASCITEEPIEHTDAIHVS 413
Qy 433 FPEFMDLMLAGLAK 446
Db 414 YPTFEHLNLSKK 427

RESULT 10
ARO_AQUAE STANDARD; PRT; 431 AA.
O67494;
30-MAY-2000 (Rel. 39, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR AQ.1536.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000744; AAC07443.1; -
CC InterPro: IPR001986; EPSP_synthase.
CC Pfam: PF00275; EPSP_synthase; 1.
CC ProDom: PD001867; EPSP_synthase; 1.
CC PROSITE: PS00104; EPSP_SYNTHASE_1; 1.
CC PROSITE: PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferrase; Complete proteome.
CC SEQUENCE 431 AA; 47793 MW; EF842512BEE41D2A CRC64;
```

Query Match 33.5%; Score 766.5; DB 1; Length 431;  
Best Local Similarity 40.7%; Pred. No. 9,7e-41;  
Matches 174; Conservative 73; Mismatches 169; Indels 11; Gaps 6;

Qy 14 KSSGLSGTVRIPEDKSIHRSFNGGLASGETRTGLLEGGEDVINTGKAMQAGARIRKE 73  
Db 6 KIKRVKGLRVPDSKSIHTRAPILGALASGETLVKRPKLTISGDTLATLELTKAIRTKVREG 65

```
Qy 74 GDTWIIDVGVNGGLLAPEALPDFGNAATGCRLTMTGLGVYDFDSTFIGDASLTKKRPMGRV 133
Db 66 KEVLEIEG-RNTELEPHDVLDKAKNSGTTARTIMSGVGLSTQPFVSULTGDSLRNRPMLRV 124
Qy 134 LNPLREMGVQVKS-EDGDRLPVTLRGKPTPTITTRVPMSAQVKSASVALLAGLINTPGITT 192
Db 125 VEPLREMGAKIDGREBGNKLPATIRGNL-KGISYFNKKSAQVKSALLAGLRABGMTE 183
Qy 193 VIEPTMTDRHTKMLQGFQANLTVETDADGVRTIRLEGRGKLTGGVIDVPGDPSSTAPPL 252
Db 184 VVEPYLSRDHTERMLKLFQAEVITTPERG-HIVKIKGQELQGTVEYCPADPSSAAYFA 242
Qy 253 VAALLVPGSDVTILANVMNPTRTGLTLTQEMGADIEVINPLAGGEDVADLRVRS-SFL 311
Db 243 ALATLAPEGEIRLKEVLLNPTROGFRKLIEMGGDISFENYRELSNEPMADLVVRVDNL 302
Qy 312 KGVTPEDRAPSMIDYPIILAVAAFAEAGATVNMGLEELRVKESDRLSAVANGKLNQVDC 371
Db 303 KPVKVSPEEVPITLIDPIILAVMAFADGVSEVKGAKELRYKESDRIKAIVTNLRKLQVQ 362
Qy 372 CDEGETSLVVRGPDCKGLGNAGAAVATHLDRHIAVMSFLVNGLVSENPVYDDATMIAT 431
Db 363 VEEFEDGFAIHGTKEIK-----GGVIETFKDHRITAMAFVLGLVVEEVIIDHPCEVTV 416
Qy 432 SPPEFMD 438
Db 417 SYPERWE 423

RESULT 11
ARO_AQUAE STANDARD; PRT; 410 AA.
AC O9WYIO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR TM0345.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MS98 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001715; AAD35431.1; -
CC TIGR: TM0345; -
```

Thu Aug 15 13:49:59 2002

DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00375; EPSP\_synthase; 1.  
DR Prodom: PD001967; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE: PS00985; EPSP\_SYNTHASE\_2; 1.  
DR Aromatic amino acid biosynthesis; Transferrase; Complete proteome.  
SQ SEQUENCE 410 AA; 45341 MW; DC9F7D474752CA69F\_CK664;

Query Match 30.4%; Score 695; DB 1; Length 410;  
Best Local Similarity 40.5%; -Pred. No. 2.5e-36;  
Matches 170; Conservative 67; Mismatches 161; Indels 22; Gaps 10;

Qy	24	IPGDKSIHRSFMEGLASSETRITLLEGEDVINTCKAMQAMGARIRKEGDTWIIDVG	83
Db	4	VPPDKSIHRAILSALAEFTSYLLNRLCLDTERTHDILEKLGTRF--EGD-WEKMKVF	60

QY 84 NGGLLAPENAPDFGNAATGGLRTMGLVGVYDFDSTFTGDSASLTKRPMGRVNLNPIRMGVQ 143

[illegible]

QY 204 EKMLQGFAGNLTVETDADGYRTIRLEGRGKLTQVIDVPGDPSTAFPLVAALLVPGSDV 263

180	ERMLK ---NLGVPVEVEGRV-LE-PATFRGFTMKVPGDLSAAAFVVLGAHFNARI	235
Db		
264	TILNVLNPNRTGLILTQEMGADI--EVINPLAGGEDVADLVRSS-TLKGVTVPEDR	320
Qy		

Db 234 TVTDVGLNPRTGLLGVMLKANLEWITEENL--EPIGTVRVETSPNLKGVVVPEHL 290

Gv 321 APSMTQFYPII AVAAFAEAGATVMNGLEELRVKESDRLSAVANGKLKGVDCEGETSLV 380

291 VPLMIDELPLVALLGVFAEGETVVRNAEELKKESDRIVLVENFKRLGVIEIEEFDGFK 350

Qy	381	VRGPDCKGLGNASGAAVATHLDHRIAMSEFLVMGLVSENPVTVDATMIATSFPFMDLM
Db	351	IVGKQSIIK-----GGSDVDEGHRAAMLFSIAGLVSEEGVDVKDHECVASFPPNYELL

RESULT 12	
AROA_CAMJE	
ID AROA_CAMJE	STANDARD;
	PRT; 428 AA.

AC	P52312; Q9PP36;
DT	01-OCT-1996 (Rel. 34, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last sequence update)

16-QCT-2001 (Rel. 40, Last annotation update)  
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
AP0A\_06310895C

OS *Campylobacter jejuni*.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC *Campylobacter*.

OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=81116;  
RX MEDLINE=97128776; PubMed=8973316;  
RA Woesten M.M.S.M., Dubbink V.H.J., van der Zeijst B.A.M.;  
PM Imbo and van der Zeijst B.A.M. 1997. Characterization of

RP SEQUENCE FROM N.A.  
RN [2]  
RL Gene 181:109-112(1996).  
RT The *aroA* gene of *Campylobacter jejuni*;

RC STRAIN=NCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,

RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,  
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrell B.G.?  
RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*

```

RT reveals hypervariable sequences." ;
RL Nature 403:665-668(2000).
CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -|- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC ----- X89371; CAA61554.1; -
DR EMBL; X89371; CAA61554.1; -
DR EMBL; AL139076; CAB73153.1; -
DR EMBL; AL139076; CAB73153.1; -
CC -----
```

```

DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase. 1.
DR ProDom: PD001867; EPSP_synthase. 1.
DR PROSITE: PS00104; EPSP_SYNTHASE. 1.

```

DR	PROSITE, P30088; EPSE-SYNTHASE.2; 1.	
DR	PROSITE, P30088; EPSE-SYNTHASE.2; 1.	
KW	Aromatic amino acid biosynthesis; Transferase; Complete proteome.	
FT	CONFLICT 152 152 K -> N (IN REF. 1).	

FT	CONFLICT	158	F -> Y (IN REF. 1).
FT	CONFLICT	179	D -> N (IN REF. 1).
FT	CONFLICT	183	T -> A (IN REF. 1).

FT	217	N -> S (IN REF. 1).
FT	217	CONFLICT
FT	241	V -> A (IN REF. 1).
FT	241	CONFLICT
FT	428	AA; 47288 MW; 465E2B50F39FC5AB CRC64;
SQ	SEQUENCE	

Query Match	25.8%	Score	590	DB 1	Length	428
Best Local Similarity	34.0%	Pred. NO.	8.5e-30			
Matches 146; Conservative		94; Mismatches	167; Indels	22; Gaps	11;	

QY 24 IPGDKSISHRSFMEGGLASGETRITGLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVG 83

[illegible]

Db 75 PKEILSPNCILDCNGSGTMRMTGFLAGISGF---FVLSGDKYLNRRPMRRISKPLTQI 131

QY 141 GVQVKSEGDRL-PVTLRGPKTPPTITYKVPMSAQVKSAVLLAGLNTPTTVEPMT 199

db 132 GARYGRNEANLAPLCIEGOKLKA-FNFKSEISSAQVKTMILSAFRADNVCTFSEISLS 190

QY 200 RDHEKMLQGFGANLTVETDADGVNRIKREGGRGLAQI LEVFOU OUCUMAKCOT  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
191 RNSENMLKAMKAPIRSVNDGSLSEINPL--KKPLKAQNIIIPNDSSAFYFVLAAIILP 248

Db

Qy 260 GSDVTILNVLMPRTGLTLTLOENGADIEVINPLAGEDVADRVRSTLKGTVTPD 319  
- - - : : : : :  
Db 249 KSOIILKNLLNPRIEAYKILQRMGAKLEMTITQ-NDETIGEIRVESKLNGIEY-KD 306

[illegible]

Db 307 NIAWLIDEAPALATAFALAKGSSLINAKSEKVRKESDRKAVVWENLQGVLRKDEEDSR 439

QY 380 VVGRPDGKGLGNASGAAVATHLDHRIAMSFVLMGLVSENPVTVDATMTATSPFPEFMDL 439

DB	367	EIEGCELK-----SSKISGDHRIAMFAILLCLL--GIEIDSDCIKTSFFNFIEI	418
OV	440	MAGIGAKIE	448

419 LSNLGARID 427

## RESULT 13

```

AROQ_STAAU          STANDARD;          PRT;          430 AA.
ID AROQ_STAAU          PRT;          430 AA.
AC Q05615;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxId=1280;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 8325-4;
RC MEDLINE=93381456; PubMed=83711108;
RA O'Connell C.M., Pattee P., Foster T.J.;
RT "Sequence and mapping of the aroA gene of Staphylococcus aureus
8325-4."
J. Gen. Microbiol. 139:1449-1460 (1993).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05004; AAA71897.1; -
CC InterPro; IPR001986; EPSP_synthase.
CC Pfam; PF00275; EPSP_synthase; 1.
CC ProDom; PD001867; EPSP_synthase; 1.
CC PROSITE; PS001104; EPSP_SYNTHASE_1; 1.
CC PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferase.
CC SEQUENCE 430 AA; 46852 MW; E5FC878EA1C23C20 CRC64;
-----
Query Match 25.2%; Score 577; DB 1; Length 430;
1st Local Similarity 33.6%; Pred. No. 5.5e-29;
Matches 144; Conservative 82; Mismatches 179; Indels 24; Gaps 8;
QY 15 SSGLSGTVRIPGDKSISHRFMGGLASGSTRITGLLEGEDVINTGKAMQAMGARIRKEG 74
DB | : : : : : | | | : : : | | | : : : | | | : : : |
10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGVSTYKPLLGDCRRTMDFPHLGVEIKEDD 69
QY 75 DTWIDGVNGGLLAPAPLDFGNAATGCRLTGMLGVGYDFDSTFIGDASLTRKPMGRVL 134
DB | : : : : : | | : : : | | : : : | | : : : | | : : : |
70 EKLWVTPGY-QVNTPHQVLYTGNSGTTRTLLAGLLSGLGNESVLSGDSVSTGKRPMDRV 128
QY 135 NPLREMGVQVKSEGDRLPVLTKGPKTPPTITRVPMASQVKSAVLLAGLNPTGITTVI 194
DB | : : : : : | : : : | : : : | : : : | : : : | : : : |
129 RPLKMDANTEGIEDNTPLIUK-PSVIKQINQMEVASAQVKSAILFASLFKEPTIIK 187
QY 195 EPLWTRDHTKMLQGF-----GANLTVETDADGVRTIRLEGRKLTQVIDVGDPSSTA 249
DB | : : : : : | | | : : : | | : : : | : : : | : : : |
188 ELDSVRNHTTMEKHFNPITEAGLSLNTTPEAIRYIKPAD-----FHPVGDITSSAA 239
QY 250 FPLVAALLVPGSDVTTILNVLNMPTRTGLITLQBMGDADIEVINPRLAGGEDVADLRV-S 308
DB | : : : : : | | : : : | | : : : | : : : | : : : | : : : |
240 FFIVAALITGSDVTHNVGINOTRSIIIDIVKMGNIQLFN-QTTGAETASIRIQYT 298
QY 309 STLKGVTVPEDRAPSMIEPYPIAAVAFAAGATVMNGLLERLKRVKLSAVANGKLN 368
DB | : : : : : | | | : : : | : : : | : : : | : : : | : : : |

```

```

Dd      299  PMLQPTTIEGELVPKAIDELPIVALLCTQAVGTSTIKDAEELKVKNRNRIDTTADMJNLL 358
Qy      369  GVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDHRIAMSFVLGVLSNPVTVDATM 428
Db      359  GFELQPTNDGLIHT--PSEFKTNATDILT---DHRIGMMLAVACVLSSEPVKIKQFDA 411
Qy      429  IATSFPEFM 437
Db      412  VNVSPFGFL 420

RESULT 14
AROA_HELPJ STANDARD;          PRT;    429 AA.
ID AROA_HELPJ Q9ZKF7;
AC AC
DT DT 30-MAY-2000 (Rel. 39, Created)
DT DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5'-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GS AROA OR JHP0980.
GN Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RR SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.I., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tunman P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESES OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: AE001527; AD06557.1; -.
DR InterPro: IPR001986; EPSP_synthase.
DR Pfam: PF00275; EPSP_synthase; 1.
DR ProDom: PD001867; EPSP_synthase; 1.
DR PROSITE: PS00104; EPSP_SYNTHASE.1; 1.
DR PROSITE: PS00885; EPSP_SYNTHASE.2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 429 AA; 47167 MW; 92724CA257241 CRC64;

Query Match                22.8%   Score 521.5; DB 1; Length 429;
Best Local Similarity     33.6%; Pred. No. 1.se-25;
Matches 145; Conservative 73; Mismatches 166; Indels 47; Gaps 13;

Qy      27  DKSTSHRSFVFGLASGETRITGLLEGEDINTCKAMQAMGARIRKEG-----DTWI 78
Db      10  DKSLSHRAVIFSLLAQPCFVRNFLMGEDCLSSLEIAQLNGAKVENTAKNSFKITPTTI 69

Qy      79  IDGVNGGLLIAEAPLPDFGNAATGCRTMTGLUGVVYDFDSTFIGDASTTKRPMGRVLNPLR 138
Db      70  KE-----PNRKILNCNNGTSWRMYLGGLSAQAGLGIVLSDGNSLNARNPKRIIEPLK 120
```

QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPPTITVYRVPMASQVKSAVLLAGLNTPGITTVIEPI 197  
D 121 AFGAKILGREDNHFAPLIVGGLPKA-CDYESPISASQVKSFAFTLSALQAQGISAYKESE 179  
QY 198 MTRDTEKMLQGFANLTVETDADGVRTIR-LEGRGKLTGQVIDVPDGPSTAPLVAAL 256  
D 180 LSRNHEIMKSLGANI---NQDGVKISPLE--KPLESFDFIANDPSSAFVALACA 234  
QY 257 LVPGSDVTILNVLNPTRTGLTLLOEMGADIE-VINPRLAGGEDVADLRVRSSTLKGVT 315  
D 235 ITPKSRLLKKNVLLNPTRIEAFVLLKMGASIEVIAQSKDL---EMIGDIIVEHAPL 292  
QY 316 VPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKNGVDCDEG 375  
D 293 IDQNTIA-SLIDEIPALSIAMLFAGKSMVKNADLRKAKESDRISKAVVSNFKALGIEEFP 351  
QY 376 ETSLVVRGRPDGKGLGNAS-----GAAVATHLDHRIAMFLVGLVSENPVTVD 425  
D 352 EDGFI-----EGLGDASQLKQHSKTKPIIKSFENDHRIAMFAVLTAL--PLEIDN 403  
D 426 ATMATSPPEF 436  
D 404 LECANISFPTE 414

RESULT 15  
AROA\_HELPY  
ID AROA\_HELPY STANDARD; PRT; 429 AA.  
AC P56197;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
GN AROA OR HP0401.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathley L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
Pylori".  
RL Nature 388:539-547(1997).  
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; AE000556; AND07470.1; .

TIGR: HP0401; .  
DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
SQ SEQUENCE 429 AA; 47240 MW; 19545753E081FDAE CRC64;

Query Match 22.7%; Score 518.5; DB 1; Length 429;  
Best Local Similarity 33.5%; Pred. No. 2.3e-25;  
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

QY 27 DKSISHRSMFPGIAGSETRITGLLEGEDVINTGKAMQAMGARIRKEG-----DTWI 78  
D 10 DKSLSHRAVIFSLAOKPCFVRNFMGDCDLSLEIAQNLGAKVENTAKNSFKITPPTTI 69  
QY 79 IDGVNGGLLAPEAPLDFGNATGCRITMGLVGVYDFDSTFIGDASLTKRPMSVRLNPLR 138  
D 70 KE-----PNKILNCNNSGTTMRLYSGLLSAQKGLFVLSGDNLSNARPMKRIIEPLK 120  
QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPPTIT---YRVPMASQVKSAVLLAGLNTPGITTVI 194  
D 121 AFGAKILGREDNHFAPLIVLG----SPLKACHYESPIASQVKSFAFTLSALQAQGASTYK 176  
QY 195 EPIMTRDHTERKMLQGFANLTVETDADGVRTIR-LEGRGKLTGQVIDVPDGPSTAPPLV 253  
D 177 ESELSRNHTEIMKSLGADI---HNQDGVKISPLE--KPLEAFDFIANDPSSAFFAL 231  
QY 254 AALVPGSDVTILNVLNPTRTGLTLLOEMGADIE--VINPRLAGGEDVADLRVRSSTL 311  
D 232 ACATTPKSRLLKKNVLLNPTRIEAFVLLKMGASIEVIAQSKDL---EMIGDIIVEHAPL 288  
QY 312 KGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKNGVD 371  
D 289 KAINIDQNTIA-SLIDEIPALSIAMLFAGKSMVKNADLRKAKESDRISKAVVSNFKALGIE 347  
QY 372 CDEGETSLVVRGRPDGKGL---GNASCAAVATHLDHRIAMFLVGLVSENPVTVDAT 427  
D 348 CEEFEDGEYVEGLEDISPLKQRFESRIKPLIKSFNDHRIAMFAVLTAL--PLEIDNLE 405  
QY 428 MATSPPEFMDLM 440  
D 406 CANISFPQFKHLL 418

Search completed: August 15, 2002, 14:06:07  
Job time: 406 sec





**THIS PAGE BLANK (USPTO)**



C:Accession: E83250  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa  
Nature 406, 959-964, 2000  
A:Reference number: A82950; MUID:20437337  
A:Accession: E83250  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-746 <STO>  
A:Cross-references: GB:AE004740; GB:AE004091; NID:g9949275; PIDN:AG06552.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3164

Query Match 43.2%; Score 989.5; DB 2; Length 746;  
Best Local Similarity 48.5%; Pred. No. 1.7e-55;  
Matches 214; Conservative 62; Mismatches 152; Indels 13; Gaps 3;

QY 12 ARKSSGLSTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71  
DB 317 AOPGSLSTGTRVPGDKSISHRSIMLSLAEGETTEVEGFELEGEDALATIAQAFDMGVVIE 376  
QY 72 -KEGDTWIIDGVNGGLLAPDLDFGNATGCRLTMTGLVGVYDFDSTFIGDASITKRP 130  
DB 377 GPQGRVTVHGVGLHGLKAPPGIYLGNSGTSMRLSLAQAQPPDSTLTGDSLSKRP 436  
QY 131 GYVLNPLREMGVQVKSDDGLPVTLRGPKTPTTYRVPMASQVKSVAVLGLAGLNTPGI 190  
DB 437 NRVAKLREMGAVIETGPEGPPPIRGQRLTGMYHMDPMASQVKSCLLLAGLYAAGE 496  
QY 191 TTIVPIPTRDHTEKMLQGFANLTVDADGVRIRLEGRGKLTGQVIDVPGDSSSTAF 250  
DB 497 TSVTEPAPTRDHTERMLRGFGYVYVVEGS-----TAKVESGHKLKSAHIEVPADISSAAF 551  
QY 251 PLVAALLVPGSDVTILNVLNMTPTGLTLTLOEMGADIEVINPRLAGGADVADLRVRSST 310  
DB 552 FLVAASIAEGSELVLOHVGINPRTVGVVIELRLMGDDLSENRQREVGGEPVADIRVRSR 611  
QY 311 LKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNV 370  
DB 612 LKSIDIPEDLVLAIDEPFVFAAACAEGRVLRGAEEELRVKESDRIOVMADGLKALGV 671  
QY 371 DDEGETSLVVRGPDGKGLGNASGAATVLDHRLHRLMFLVGLVSENPVTVDDATMIA 430  
DB 672 KAEPTPDGIVIEG-----GAFGGGEVWAHGDHRIAMSFVSASLRASGPRIHDCANVA 724  
QY 431 TSPEFMDLMAGLGAKEILSD 451  
DB 725 TSPNFLALCAQTGIRVAVEN 745

RESULT 4  
A82572  
3-phosphoshikimate 1-carboxyvinyltransferase XF2324 [imported] - Xylella fastidiosa (C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: A82572  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: A82572  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-454 <SIM>  
A:Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85123.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer

QY 361 VANGKLKLVGDCDEGETSLVVRGPDGKGLGNASGAATVLDHRLHRLMFLVGLVSENP 420  
DB 391 VARGLEANGVDCTEGEMSLTVRGPGRGLG---GGTVGTHDLHRIAMSFVLMGLASEKP 447  
QY 421 VTVDATMIATISFPFMDLMAGLGAKI 447  
DB 448 VTVDSTMIATISFPFMDLMAGLGAKI 474

RESULT 2  
C87694  
3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Caulobacter crescentus (C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: C87694  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Emdolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: C87694  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-443 <STO>  
A:Cross-references: GB:AE005673; NID:g13425333; PIDN:AAK25551.1; GSPDB:GN00148  
C:Genetics:  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy  
Query Match 49.6%; Score 1135; DB 2; Length 443;  
Best Local Similarity 55.8%; Pred. No. 4.4e-65;  
Matches 251; Conservative 45; Mismatches 138; Indels 16; Gaps 6;

QY 2 LHGASSRPATARKSSGTSVTRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGK 61  
DB 3 LAGLKSAPGGA-----LRGIVRAPGDKSISHRSMILGALATGTTTVEGLEGGDVLTAR 57  
QY 62 AMQAMGARIRKEG-DTWIIDGVNGGLLAPDLDFGNATGCRLTMTGLVGVYDFDSTFI 120  
DB 58 AMQAFGARIRREGVGRWRIE--KGKGFEEVPVVIDCGNAGTVGRLLMGAAAGFAMCATFT 115  
QY 121 GDASLTARPMGRVNLREMGVQVKSDDGLPVTLRGPKTPTTYRVPMASQVKSVA 180  
DB 116 GDQSLRGRPMGRVLDPLARMGATWLGKRLPLTLKGGNL-RGLNVTLPMSAQVKSVA 174  
QY 181 LLAGLNTPGITTVIEPIWTRDHTKMLQGFANLTVDADGVRT---IRLEGRGKLTGQ 237  
DB 175 LLAGLHAEGGVETIEPEATRDHTERMLRAFGEVIEDRKAGDKTFRHLRPEGQKLTGT 234  
QY 238 VIDVPGDPSSAFPLVAALLVPGSDVTILNVLNMTPTGLTLTLOEMGADIEVINPRLAG 297  
DB 235 HVAVPGDPSSAAFPVLAALVPGSEVTEGVMLNELRTGLFTLLOEMGADLVISNVRAS 294  
QY 298 GEDVADLRVRSSTLKVTVTPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDR 357  
DB 295 GEEVGDTIARYSQLKGVVPPERAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDR 354  
QY 358 LSAVANGKLKLVGDCDEGETSLVVRGPDGKGLGNASGAATVLDHRLHRLMFLVGLV 417  
DB 355 ISLTANGLKACGVQVVEPEGFIV---TGTGQPPKGGATVTVTHGDHRIAMSHUILGMAA 410  
QY 418 ENPVTVDATMIATISFPFMDLMAGLGAKI 447  
DB 411 QAEVAVDEPGMIATISFPFMDLMAGLGATL 440

RESULT 3  
E83250  
3-phosphoshikimate 1-CARBOXYVINYLTRANSFERASE prephenate dehydrogenase H  
still frameshift 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE prephenate dehydrogenase H  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigh  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2324  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox  
Query Match 39.9%; Score 912.5; DB 2; Length 454;  
Best Local Similarity 46.9%; Pred. No. 7.1e-51;  
Matches 206; Conservative 60; Mismatches 158; Indels 15; Gaps 4;  
QY 12 ARKSSGLSTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71  
Db 23 AHQGTPLHGVLSIPGDKSIHRAVMPAALADGTSRIDGFLEABDTCTAEILARLGVRIE 82  
QY 72 KEGDTW-IIDGVNGGLLAPEAPLDFGNATGCRLTMLGVGVYDFDSTFIGDASLTKRPM 130  
Db 83 TPLSTQRIHVHGVVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPPDSVLVGDASLTKRPM 142  
QY 131 GRVLPNLEMGVQVKSDEGDRPLVTLRGPKTPPIRYRPMASQVKSAAVLLAGLTPGI 190  
Db 143 RVTDPLOMGARIDSDGTPPLRIYQGLLHGIDFISPVASAQKSAVLLAGLVARNE 202  
QY 191 TVVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGOVIVDPGDPSPSTAF 250  
Db 203 TVVREPHPTDYTERMTALFGVDIDVSTGC-----ARLRGGORLCATDITIPADFSAAF 257  
QY 251 PLVAALLVPGSDVTILNVLNMPRTGLILTLOEMGADIEVINPRLAGGEDVADLRVRSST 310  
Db 258 YLVAASVPGSDILTIRAVGNFRIGLLTVLRMGANIVESNRHGOEGEPVVDLRVRYAP 317  
QY 311 LKGVTVPEDRAPSMIDEPYPLAVAAFAFAGATVYMGLEELRVKESDRLSAVANGKLNGV 370  
Db 318 LQCTRVPELVDADMIDEPFALFVAAAEEQVTVSAAELRVKESDRLAAMVTGLRVLG 377  
QY 371 DCEGETSLVVRGPRDCKGLGNASGAATHLDHRTAMFLVGLVSENPVTVDATMTIA 430  
Db 378 QVDETADGATIRGGPTGCHTINSHG-----DHRTAMAFSIAGQLSVSTRIEDVANVA 430  
431 TSFPEFMDL--MAGLGAKI 447  
Db 431 TSFPDYETLARSAGFLEV 449  
RESULT 5  
5-enolpyruvylshikimate-3-phosphate synthase aroE [imported] - Bacillus halodurans (stra  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: C83858  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: C83858  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-431 <STO>  
A:Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB05386.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: aroE  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

Query Match 38.6%; Score 884; DB 2; Length 431;  
Best Local Similarity 44.2%; Pred. No. 4.4e-49;  
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;  
QY 15 SSGLSGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74  
Db 10 AKGLKTIKVPDGKSIHRAVMPALAKGTTTVEGFLPGADCLSTISCFKGLGVSTEQAE 69  
QY 75 DTWIIDGVNGGLLAPEAPLDFGNATGCRLTMLGVGVYDFDSTFIGDASLTKRPMGRVL 134  
Db 70 ERYTVKKGWGLREPSDILDVNGSGTTRTLILGILSTLFPFHSVIGDESIGRPMKRV 129  
QY 135 NPUREMGVQVKSDE-GDRPLVTLRGPKTPPIRYRPMASQVKSAAVLLAGLTPGTITV 193  
Db 130 EPLKMGAAQIDGRDHGNLTPLSIRGGOL-KGIDFHSVSAQMKSAIILLAGLRAEGKTSV 188  
QY 194 IEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGOVIVDPGDPSPSTAFPLV 253  
Db 189 TEPAKTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGOVIVDPGDPSPSTAFPLV 243  
QY 254 AALLVPGSDVTILNVLNMPRTGLILTLOEMGADIEVINPRLAGGEDVADLRVRSSTPLK 313  
Db 244 AGAMVPHSRITLNVGINPTRAGILEVLKQMGATLAMENRVOGGEVADLTITSVLQ 303  
QY 314 VTVPEDRAPSMIDEPYPLAVAAFAFAGATVYMGLEELRVKESDRLSAVANGKLNGVDCD 373  
Db 304 VEIIGDIIPLRIDEIPIIAVLAQAGRTVTKDAEELKVKETNRIDTVSELKLGASIH 363  
QY 374 EGETSLVVRGPRDCKGLGNASGAATHLDHRTAMFLVGLVSENPVTVDATMTATSF 433  
Db 364 ATDDGMIIEGPTPLK-----GVTVSSHGDRHIGMAMIAAALLAEKXPVTVEGTEATAVS 418  
QY 434 PEPM 438  
Db 419 PSFFD 423  
RESULT 6  
S76218  
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Synechocystis sp. (stra  
N:Alternate names: 5-enolpyruvylshikimate 3-phosphate synthase  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S76218; S33748; S34393  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76218  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-447 <KAN>  
A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BAAL8477.1; PID:g165  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996  
R:Mayes, S.R.; dalla Chiesa, M.; Zhang, Z.; Barber, J.  
FEBS Lett. 325, 255-261, 1993  
A:Title: The genes aroA and trnQ are located upstream of psbO in the chromosome of Sy  
A:Reference number: S33748; MUID:93307506  
A:Accession: S33748  
A:Molecule type: DNA  
A:Residues: 103-194 <MAY>  
A:Cross-references: EMBL:X72784; NID:g394747; PIDN:CAA51291.1; PID:g394748  
C:Genetics:  
A:Gene: aroA  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car  
C:Keywords: aromatic amino acid biosynthesis; transferase  
F:28-439/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>



Db 296 QLKGTGTEISDLPLRIDELPIIALATOGQOTIIRDAELKVKEDRIVAVANALNSMG 355  
QY 370 VDCDEGETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLMGLVSEN-PVTVDATM 428  
Db 356 AKIQPTDDGMIIOG---GPKL-HAPENSINTGLDHRIGMAAIAALLVKNGEIELERA 411  
QY 429 IATSPFPEMDLMAGLAKI 447  
Db 412 IQTSYPSFDDLEQLSENI 430

RESULT 9  
S44096  
enolpyruvylshikimate 3-phosphate synthase - Dichelobacter nodosus  
C:Species: Dichelobacter nodosus  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
A:Accession: S44096  
R:Alm, R.A.; Dalrymple, B.P.; Mattick, J.S.  
Description: Sequencing and expression of the *aroA* gene from *Dichelobacter nodosus*.  
Reference number: S44096  
A:Accession: S44096  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-443 <ALM>  
A:Cross-references: EMBL:Z29339; NID:g472931; PIDN:CAA82544.1; PID:g472932  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy F:16-424/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

Query Match 35.2%; Score 806; DB 2; Length 443;  
Best Local Similarity 42.2%; Pred. No. 4.3e-44;  
Matches 185; Conservative 74; Mismatches 169; Indels 10; Gaps 3;

QY 16 SGLSTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIRKEGD 75  
Db 12 SALSGETITCGDKSNHRALLAALAEQTEIRGFLACADCLATRALGVDIQREKE 71  
QY 76 TWIDGVNGGLLAEAPLDFGNATGCRLTGMLGVGVYDFDSTFTIGDASLTKRPMGRVLN 135  
Db 72 IVTIRGVGLGLOPKAPLNMQNSGTSMLLAGILAAQAFESVLCGDESLEKRRPMQRIIT 131  
QY 136 PLREMGVOVKSEDDRLPVTLRGPKTPTTYRVPMAAQKSAVLLAGLNTPGIITVIE 195  
Db 132 PLVQMGAKIVSHNSTAPLHISG-RPLTGIDYALPLPSAQLASCLILAGLLADGTYRLHT 190  
QY 196 PLWTRDTEKMLQGFANLTVETDADGVRTIRLEGRKGLTGQVIDVPGDPSSTAPFLVAA 255  
191 CGISRDRHTERMLPFGGALEIKKE-----QIIIVTGGQKLHGCVLDIVGDLSSAAAPFMVAA 245  
QY 256 LLVPGSDVTILNVLNPTFTGLILILQENGADIEVINPRLAGEDVADLRVRSSTLKGVT 315  
Db 246 LIAPRAEVVIRNVPNTPTAAITLLQRMGRIELHQHGFWEAEPVADVIVYHSLKRGIT 305  
QY 316 VPEDRAPSMIDRYPILAVAAFAEGATVNGLEELRVKESDRLSAVANGKLNGVDCDEG 375  
Db 306 VAPETIANADELPFFIAACAEGETTFVGNLSSELVKESDRLSAAQNQLTIGVACDVG 365  
QY 376 ETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLMGLVSENPTVDDATMIATSPFE 435  
Db 366 ADFIHYGRSDRQFL----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMPO 421  
QY 436 FMDLMAGLCAKLELSDTK 453  
Db 422 FRDFAAAGMNVGEKDAK 439

RESULT 10  
D95159  
3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Streptococcus pneumoniae (stra  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C:Accession: D95159  
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris  
A:Title: Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: D95159  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-427 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK75469.1; PID:g14972856; GSPDB:GN00164; TIGR:  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP1371  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car

Query Match 35.1%; Score 803.5; DB 2; Length 427;  
Best Local Similarity 42.8%; Pred. No. 5.9e-44;  
Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIRKEGDTW 77  
Db 10 LHGSIRVPDGKSKSHRSIIIFGSLAEGTKVYDILRGEDVLSTMQVFRDLGVEIEDKQVI 69  
QY 78 IIDGVNGGLLAEAPLDFGNATGCRLTGMLGVGVYDFDSTFTIGDASLTKRPMGRVLNPL 137  
Db 70 TIQGVGMAGLKAPONALNMNGSGTSIRLISGLVLAGADFEVEMFGDSDLSKRPMDRVTLPL 129  
QY 138 REMGVQVKSE-DGDRLPVLIRGPKTPTTYRVPMAAQKSAVLLAGLNTPGIITVIEP 196  
Db 130 KKMVSISQOTERDLPLRLKGTNKLRIHYELPIASAQVKSALMFAALQAQGESVIEK 189  
QY 197 IMTRDTEKMLQGFANLTVETDADGVRTIRLEGRKGLTGQVIDVPGDPSSTAPFLVAAAL 256  
Db 190 EYTRNTEMDLKFQGGHLSV----DG-KKITVQGPQKLTGQKVVVFGDLSAAFWLAVGL 244  
QY 257 LVPGSVDTILNVLNPTFTGLILTLQEMGADIEV--INPRLAGEDVADLRVRSSTLKG 314  
Db 245 IAPNSRLVLQNVGINETRTGIIDVIRAMGGKLEITEIDPVAKS---ATLIVESSDLKGT 300  
QY 315 TVPEDRAPSMIDRYPILAVAAFAEGATVNGLEELRVKESDRLSAVANGKLNGVDCDE 374  
Db 301 EIGGALIPRLIDELPIIALATOAQGVTVIKDAEELVKRETDRIQVVDALNSMGADITP 360  
QY 375 GETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLMG-LVSENPVTVDDATMIATSF 433  
Db 361 TADGMIKCK-----SALHGARVNTFGDHRIGMMTAIAALLVADGEVELDRAEINTSY 414  
QY 434 PFMMDLMAGL 443  
Db 415 PSFDDLES 424

RESULT 11  
S52580  
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Lactococcus lactis  
N:Alternate names: 5-enolpyruvylshikimate-3-phosphate synthase  
C:Species: Lactococcus lactis  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
A:Accession: S52580  
R:Griffin, H.G.; Gasson, M.J.  
Mol. Gen. Genet. 246, 119-127, 1995  
A:Title: Genetic aspects of aromatic amino acid biosynthesis in *Lactococcus lactis*.  
A:Reference number: S52579; MUID:95124293  
A:Accession: S52580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-430 <GRI>  
A:Cross-references: EMBL:X78413; NID:g683581; PIDN:CAA55180.1; PID:g683583  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car

C:Keywords: transference  
F:12-421/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

Query Match 35.18; Score 802.5; DB 2; Length 430;  
Best Local Similarity 41.28; Pred. No. 6.9e-44;  
Matches 181; Conservative 79; Mismatches 158; Indels 21; Gaps 7;

QY 15 SSGLSGTVRIIPGDKSISHRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEG 74  
DB 7 SGLAGRLKVPDKSISHRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEG 66

QY 75 DWIIIDGVNGGLLAPEAPLDFGNAATGCRITGLVGVYDFDSIFIGDASLTGKPMGRVL 134  
DB 67 QVITVHGOGISKLKEPEKALONGSGTSTRLLSGLIAGLPEATLFGDSSLSKRPMDRVA 126

QY 135 NPLREMGVQVKE-DGDRLPVTLRGPKTPTTYRVPMSAQVKSAYVLLAGLNTGPIITV 193  
DB 127 TPLQMGAGIIVQOTDKVLPMTIKGSAHLKADYILPVASQVKSAYVLLAGLNTGPIITV 186

QY 194 IEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253  
DB 187 VEKXTRSHTEMLVQFGGEITVSD-----KTILVPGGOKLLGQEVTPVPGDISSAAFLV 241

QY 254 AALLVPGSDVTILNVLNMPRTGLIILTLQEMGADIEVINPRLAGGED---VADLRVRSS 309  
DB 242 AGLWVNSGLILENVTGILVETGILVETGILVETGILVETGILVETGILVETGILVET 295

QY 310 TLKGVTPEDRAPSMIDEXPIILAVAAFAEGATVMNGLEELRVKESDRLSAYVANGKLNG 369  
DB 296 QLKGTISGLDLPRLIDELPIIALLATQAGTKTIIRDAELKVKETDRIAVVADALNSMG 355

QY 370 VDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHIAFSFLVMG-LVSENPTVDDATM 428  
DB 356 ANTEPTDDGMIIQO---GTKL-HAPENAIITLGDHRIGMVAIAALLVNGEIELEERA 411

QY 429 IATSEPEFMDLMAGLGAKI 447  
DB 412 IQTYSFDFDLEKLSGNL 430

RESULT 12  
D98025  
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) [imported] - Streptococcus pneumoniae  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
C:Accession: D98025  
R: Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; Dehoff, B.S.; R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McNahren, S.; P.; Sun, P.M.; Winkler, M.E.  
P. Bacteriol 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: D98025  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-431 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAL00033.1; PID:g15458865; GSPDB:GN00174  
C:Genetics:  
C:Gene: aroA  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxyvinyltransferase

Query Match 35.08; Score 801.5; DB 2; Length 431;  
Best Local Similarity 42.8; Pred. No. 8e-44;  
Matches 184; Conservative 72; Mismatches 155; Indels 19; Gaps 7;

QY 18 LSGTVRIIPGDKSISHRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEGDTW 77  
DB 14 LHGSIRVPGDKSISHRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEGDTW 73

QY 78 IIDGVNGGLLAPEAPLDFGNAATGCRITGLVGVYDFDSTFTIGDASLTGKPMGRVNLPL 137  
DB 74 TIQGVNGAGLKAPQALNMGNSGTSIRLSGLVAGADFEVEMFGDDSLSRPMDRVTPL 133

QY 138 REMGVQVKE-DGDRLPVTLRGPKTPTTYRVPMSAQVKSAYVLLAGLNTGPIITVIEP 196  
DB 134 KMGVSISSGQTERDLPRLKGTKNLRPIHYELPIASQVKSALMFAALQAKGESVIEK 193

QY 197 IMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256  
DB 194 ECTNRHEDMLKQEGHLSV---DG-KKIVQGPQKLTGQVVVPGDISSAAMFVAGL 248

QY 257 LVPQSDVTILNVLNMPRTGLIILTLQEMGADIEV-INPRLAGGEDVADLRVRSSTLKV 314  
DB 249 INPNSHLVQVNGINETGTIIVIRMGKLEVIDPVAKS-----STLTVESDDLKGT 304

QY 315 TVPEDRAPSMIDEXPIILAVAAFAEGATVMNGLEELRVKESDRLSAYVANGKLNGVDCDE 374  
DB 305 BIGGALPRLIDELPIIALLATQAGTKTIIRDAELKVKETDRIAVVADALNSMGADITP 364

QY 375 GETSLVVRGPDGKGLGNASGAAVATHLDRHIAFSFLVMG-LVSENPTVDDATMIATSF 433  
DB 365 TADGMINKG-----SALHGARVNTFGDHRIGMVAIAALLVADGEVELDRAEAINTSY 418

QY 434 PEFMDLMAGL 443  
DB 419 PSFEDDLES 428

RESULT 13  
AC1315  
5-encolpyruvylshikimate-3-phosphate synthase homolog aroE [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AC1315  
R: Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla. A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1315  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-428 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00001.1; PID:g16411376; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
C:Gene: aroE  
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxyvinyltransferase

Query Match 34.9%; Score 799.5; DB 2; Length 428;  
Best Local Similarity 39.9%; Pred. No. 1.1e-43;  
Matches 169; Conservative 84; Mismatches 158; Indels 13; Gaps 5;

QY 17 LSGTVRIIPGDKSISHRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEGDT 76  
DB 9 GLVGATVTPGDKSMHSRSMFGLASGETRITGLLEGEVDINTGKAMQAMGARIRKEGDT 68

QY 77 WIIDGVNGGLLAPEAPLDFGNAATGCRITGLVGVYDFDSTFTIGDASLTGKPMGRVNLPL 136  
DB 69 IVHGVGFDGLKQADGPDLDIGNSGTTIRLMNGILAGRDFDTVLGDESIAKRPNNRMLP 128

QY 137 LREMGVQVKE-DGDRLPVTLRGPKTPTTYRVPMSAQVKSAYVLLAGLNTGPIITVIE 195  
DB 129 LQMGAKMKHKGDCSEFAPITITGKSLKRMVHMPVASQVKSATIAFAALQAKGESVIE 188

QY 196 PIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVA 255  
DB 189 KEKTRDHTMIRQFGG-----EIMDGL-TIRVKGOKFTGQEMTVPGDVSSAAFTVAG 243





Thu Aug 15 13:49:58 2002

```

| | : |::||| |||| |::| : : | :||| ||||| : | : | |
Db 303 KPVKSPEEPTLIDEIPILAVIMAFADGVSEVKAGKELRYKESDRKAIIVTNLRKLGVO 362
QY 372 CDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFVLMGLVSENPVTVDDATMIAT 431
Db 363 VEEFEDGFAIHGTKEIK-----GGVIETFKDHRIAMAFVGLVVEEVIIDHPECVTV 416
QY 432 SPPEFMD 438
Db 417 SYPEFEW 423
```

Search completed: August 15, 2002, 13:59:57  
Job time: 131 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:16 ; Search time 22.39 Seconds  
(without alignments)  
496.366 Million cell updates/sec

Title: US-09-464-099A-70

Perfect score: 2288

Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2282	99.7	455	1	US-08-476-008-3
2	2282	99.7	455	1	US-08-306-063-3
3	2282	99.7	455	1	US-08-833-485-3
4	2282	99.7	455	4	US-09-137-440-3
5	2282	99.7	455	5	PCT-US91-06148A-3
6	1900.5	83.1	449	1	US-08-476-008-5
7	1900.5	83.1	449	1	US-08-476-008-7
8	1900.5	83.1	449	1	US-08-306-063-5
9	1900.5	83.1	449	1	US-08-306-063-7
10	1900.5	83.1	449	1	US-08-833-485-5
11	1900.5	83.1	449	1	US-08-833-485-7
12	1900.5	83.1	449	4	US-09-137-440-5
13	1900.5	83.1	449	4	US-09-137-440-7
14	1900.5	83.1	449	5	PCT-US91-06148A-5
15	1900.5	83.1	449	5	PCT-US91-06148A-7
16	867.5	37.9	447	1	US-08-476-008-67
17	867.5	37.9	447	1	US-08-306-063-67
18	867.5	37.9	447	1	US-08-833-485-67
19	867.5	37.9	447	4	US-09-137-440-67
20	806	35.2	443	1	US-08-476-008-69
21	806	35.2	443	1	US-08-306-063-69
22	806	35.2	443	1	US-08-833-485-69
23	806	35.2	443	4	US-09-137-440-69
24	803.5	35.1	427	4	US-09-325-881-2
25	798.5	34.9	427	2	US-08-896-345-2
26	798.5	34.9	427	2	US-09-226-091-2
27	785	34.3	428	1	US-08-476-008-42

28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appl
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appl
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appl
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appl
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

## ALIGNMENTS

RESULT 1  
US-08-476-008-3  
Sequence 3, Application US/08476008  
Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glycosylase Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(106660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-476-008-3

Query Match 99.7%; Score 2282; DB 1; Length 455;  
Best Local Similarity 99.8%; Pred. No. 1.3e-202;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSHGASSRPATARKSSGLSGTVIRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
1 MSHGASSRPATARKSSGLSGTVIRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGNAATGCRLTMLGVYDFDSTFI 120  
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGNAATGCRLTMLGVYDFDSTFI 120  
121 GDASLTTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTITYRVPMSAQVKSAY 180  
121 GDASLTTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTITYRVPMSAQVKSAY 180  
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVEDADGVRTIRLEGRKLTGQVID 240  
181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVEDADGVRTIRLEGRKLTGQVID 240  
241 VPGDPSSTAPPLVAALLVPGSDVTILNVLMPNTRTGLTLTQEMGADIEVINPRLAGGED 300  
241 VPGDPSSTAPPLVAALLVPGSDVTILNVLMPNTRTGLTLTQEMGADIEVINPRLAGGED 300  
301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSENP 420  
421 VTVDATMIATSPPEFMDLMAGLAKIELSDTKAA 455  
421 VTVDATMIATSPPEFMDLMAGLAKIELSDTKAA 455

RESULT 2  
US-08-306-063-3  
Sequence 3, Application US/08306063  
Patent No. 5633435  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glycosylated Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991

STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-833-485-3

Query Match 99.7%; Score 2282; DB 1; Length 455;  
Best Local Similarity 99.8%; Pred. No. 1.3e-202;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHSRFSFPGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHSRFSFPGGLASGETRITGLLEGEDVINTG 60  
61 KAMQAMGARIRKRGDWTIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120  
Db 61 KAMQAMGARIRKRGDWTIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVOVKSDEGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180  
Db 121 GDASLTKRPMGRVNLPLREMGVOVKSDEGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180  
QY 181 LLAGLNTPGITTVIEPIMRDHTKMLQCGFANLTVETDADGVRTIRLEGRKLTQVID 240  
Db 181 LLAGLNTPGITTVIEPIMRDHTKMLQCGFANLTVETDADGVRTIRLEGRKLTQVID 240  
QY 241 VPGDPSSTAFFPLVAALLVPGSDVTILNVLNPNRTTGLITLQEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAFFPLVAALLVPGSDVTILNVLNPNRTTGLITLQEMGADIEVINPRLAGGED 300  
QY 301 VADLRVRSSTLKGVTVPDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
Db 301 VADLRVRSSTLKGVTVPDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
QY 361 VANGLKNGVDCDEGETSLVVRGRPGKGLGNASGAAYATHLDHRAMSLFVNGLVSENP 420  
Db 361 VANGLKNGVDCDEGETSLVVRGRPGKGLGNASGAAYATHLDHRAMSLFVNGLVSENP 420  
QY 421 VTVDDATMIATSPPEFMDLMAGLAKIELSDTKAA 455

Db 421 VTVDDATMIATSPPEFMDLMAGLAKIELSDTKAA 455  
RESULT 4  
US-09-137-440-3  
Sequence 3, Application US/09137440  
Patent No. 6248876  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/137,440  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/833,485  
FILING DATE: 07-APR-1997  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-137-440-3

Query Match 99.7%; Score 2282; DB 4; Length 455;  
Best Local Similarity 99.8%; Pred. No. 1.3e-202;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHSRFSFPGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHSRFSFPGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKRGDWTIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120  
Db 61 KAMQAMGARIRKRGDWTIIDGVNGGGLLAPEAPLDFGNATGCRLTWGLVGVYDFDSTFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVOVKSDEGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180  
Db 121 GDASLTKRPMGRVNLPLREMGVOVKSDEGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180

Db 121 GDSLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTPTTYRVPMSAQVKSAY 180  
Qy 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGCANLTVETDADGVRTIRLEGRKLTGOVID 240  
Db 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGCANLTVETDADGVRTIRLEGRKLTGOVID 240  
Qy 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300  
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDRLSA 360  
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDRLSA 360  
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGGKGLGNASGAAVATHLDRHRIAMSFVLMGLVSEN 420  
Db 361 VANGKLNGVDCDEGETSLVVRGPRDGGKGLGNASGAAVATHLDRHRIAMSFVLMGLVSEN 420  
Qy 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455  
Db 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455

RESULT 5  
PCT-US91-06148A-3  
; Sequence 3, Application PC/TUS9106148A  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
; STREET: 700 Chesterfield Village Parkway  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/06148A  
; FILING DATE: 19910828  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/749,611  
; FILING DATE: 28-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(10535)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-6099  
; TELEFAX: (314)537-6047  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 455 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US91-06148A-3

Query Match 99.7%; Score 2282; DB 5; Length 455;  
Best Local Similarity 99.8%; Pred. No. 1.3e-202;

Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
Qy 61 KAMQAMGARIRKEGDWTIIDGVNGGLLAPLDEGNAATGCRMTMGLVGVYDFDSTFI 120  
Db 61 KAMQAMGARIRKEGDWTIIDGVNGGLLAPLDEGNAATGCRMTMGLVGVYDFDSTFI 120  
Qy 121 GDASLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTPTTYRVPMSAQVKSAY 180  
Db 121 GDASLTRKPMGRVNLPLREMGVQKSEGDRLPVTLRGPKTPTPTTYRVPMSAQVKSAY 180  
Qy 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGCANLTVETDADGVRTIRLEGRKLTGOVID 240  
Db 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGCANLTVETDADGVRTIRLEGRKLTGOVID 240  
Qy 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPTRTGLTLTLOEMGADIEVINPRLAGGED 300  
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDRLSA 360  
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVMNGLEELRVKESDRLSA 360  
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGGKGLGNASGAAVATHLDRHRIAMSFVLMGLVSEN 420  
Db 361 VANGKLNGVDCDEGETSLVVRGPRDGGKGLGNASGAAVATHLDRHRIAMSFVLMGLVSEN 420  
Qy 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455  
Db 421 VTVDATMIATSFPEFMDLMAGLCAKIELSDTKAA 455

RESULT 6  
US-08-476-008-5  
; Sequence 5, Application US/08476008  
; Patent No. 5627061  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
; STREET: 700 Chesterfield Village Parkway  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,008  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/306,063  
; FILING DATE: 13-SEP-1994  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; CLASSIFICATION: 435





Thu Aug 15 13:49:55 2002

Sequence 5, Application US/08306063  
Patent No. 5633435  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306.063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-306-063-5  
Query Match 83.1%; Score 1900.5; DB 1; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASRRPATARSSGLSTVRIPOGKSIHSRSMFPGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHSASPKPATARRSEALTGEIRIPGKSIHSRSMFPGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAWQAMGARTKRGDVTWIDGVNGGLLAPEALDFGNAATGRLTMGLVGVYDFDSTFI 120  
Db 61 KAWQAMGARTKRGDVTWIDGVNGGLLAPEALDFGNAATGRLTMGLVGVYDFDSTFI 120  
QY 121 CDASLTKRPMGRVNLPLREMGVQKVEDGRLPVTILRGKPTPTTYRVPMSAQVKSVA 180  
Db 121 CDASLTKRPMGRVNLPLREMGVQKVEDGRLPVTILRGKPTPTTYRVPMSAQVKSVA 180  
QY 181 LLAGLNTPGTITVIEPTMTDHTKMLQGFGANLTVETDADGVRTIRLEGRGLTGOVID 240  
Db 181 LLAGLNTPGTITVIEPTMTDHTKMLQGFGANLTVETDADGVRTIRLEGRGLTGOVID 240  
QY 241 VPGDPSSTAPLVAALLVPGSDVTILNVLNPNRTTGLILTLQEMGADIEVINPLAGGED 300  
Db 241 VPGDPSSTAPLVAALLVPGSDVTILNVLNPNRTTGLILTLQEMGADIEVINPLAGGED 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLELRVKESDRLSA 360  
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLELRVKESDRLSA 360  
QY 361 VANGKLKNGVDCDEGETSLVVRGPPGKGLGNASGAATVATLHLDHRIAMSLVNGLYSENP 420  
Db 361 VANGKLKNGVDCDEGETSLVVRGPPGKGLGNASGAATVATLHLDHRIAMSLVNGLYSENP 420  
QY 421 VTVDATMTATSPPEFMDLMAGLCAKIELS 450  
Db 421 VTVDATMTATSPPEFMDLMAGLCAKIELS 450  
QY 418 VTDDSNMIAATSPPEFMDMMPGLCAKIELS 447  
Db 418 VTDDSNMIAATSPPEFMDMMPGLCAKIELS 447  
RESULT 9  
US-08-306-063-7  
Sequence 7, Application US/08306063  
Patent No. 5633435  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306.063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-306-063-7

Query Match 83.1%; Score 1900.5; DB 1; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASRRPATARSSGLSTVRIPOGKSIHSRSMFPGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHSASPKPATARRSEALTGEIRIPGKSIHSRSMFPGGLASGETRITGLLEGEDVINTG 60

```
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPAPLDGNAATGRLTMGLVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKEGDTWIIDGVNGGGLLQPEAALDFGNAGTGLTMGLVGVYDMKTSFI 120
QY 121 GDAASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPIYRVPMASAOQVKS 180
Db 121 GDAASLTKRPMGRVNLPLREMGVQVKAADGDRMPLTLTGPKTANPIYRVPMASAOQVKS 180
QY 181 LLAGLNTPGTITVIEPMDRTHTEKMLQGFANLTVETDADGVRTIRLEGRGLTGQVID 240
Db 181 LLAGLNTPGTITVIEPMDRTHTEKMLQGFADLTVETDKGVRHIRTGQGLVGQID 240
QY 241 VPGDPSSTAPLVAALLVPGSDVTILNLMNPTRTGLITLQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAPLVAALLVEGSDVTIRNLMNPTRTGLITLQEMGADIEVINARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATHLDRHRIAMSLVLMGLVSEN 420
Db 361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDRHRIAMSLVLMGLAAEKP 417
QY 421 VTVDATMIATSPFPEMDLMAGLAKIELS 450
Db 418 VTVDNSMIATSPFPEMDMMPGLAKIELS 447

RESULT 10
US-08-833-485-5
; Sequence 5, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
```

```
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-485-5

Query Match 83.1%; Score 1900.5; DB 1; Length 449;
Best Local Similarity 82.9%; Pred. No. 2.2e-167;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASSRPATARKSSGLSGTIVRIPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGGLLAPAPLDGNAATGRLTMGLVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKEGDTWIIDGVNGGGLLQPEAALDFGNAGTGLTMGLVGVYDMKTSFI 120
QY 121 GDAASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTLRGPKTPTPIYRVPMASAOQVKS 180
Db 121 GDAASLTKRPMGRVNLPLREMGVQVKAADGDRMPLTLTGPKTANPIYRVPMASAOQVKS 180
QY 181 LLAGLNTPGTITVIEPMDRTHTEKMLQGFANLTVETDADGVRTIRLEGRGLTGQVID 240
Db 181 LLAGLNTPGTITVIEPMDRTHTEKMLQGFADLTVETDKGVRHIRTGQGLVGQID 240
QY 241 VPGDPSSTAPLVAALLVPGSDVTILNLMNPTRTGLITLQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAPLVAALLVEGSDVTIRNLMNPTRTGLITLQEMGADIEVINARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGCATVMNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATHLDRHRIAMSLVLMGLVSEN 420
Db 361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDRHRIAMSLVLMGLAAEKP 417
QY 421 VTVDATMIATSPFPEMDLMAGLAKIELS 450
Db 418 VTVDNSMIATSPFPEMDMMPGLAKIELS 447
```

```
RESULT 11
US-08-833-485-7
; Sequence 7, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-833-485-7

Query Match 83.1%; Score 1900.5; DB 1; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASSRATARKSGSLGTVIRIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHASPKPATARRSALGTEIRIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGGLLAPLDFGNATGCRLTMLGVGVDFDSTFI 120  
Db 61 RAMQAMGARIRKEDGTWIIDGVNGGGLLAPLDFGNATGCRLTMLGVGVDFDSTFI 120  
QY 121 GDASLKRPMGRVNLPLREMGVQVKSEDDGRLPVTLLRGPKTPTPIYRVPMSAQVKSAY 180  
Db 121 GDASLKRPMGRVNLPLREMGVQVKSEDDGRLPVTLLRGPKTPTPIYRVPMSAQVKSAY 180  
QY 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240  
Db 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTTGLILTLQEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTTGLILTLQEMGADIEVINPRLAGGED 300  
QY 301 VADLRVRSSTLKGVTPEDRAPSMIDEPYILAVAAAFPAEGATVNGLEELRVKESDRLSA 360  
Db 301 VADLRVRSSTLKGVTPEDRAPSMIDEPYILAVAAAFPAEGATVNGLEELRVKESDRLSA 360  
QY 361 VANGKLKNGVDCDGETSLVVRGPDGKGLGNASGAAYATHLDRHIAAMSLVWGLVSEN 420  
Db 361 VANGKLKNGVDCDGETSLVVRGPDGKGLGNASGAAYATHLDRHIAAMSLVWGLVSEN 420  
QY 421 VTVDATMIATSPFEFMDLWGLGAKIELS 450  
Db 421 VTVDATMIATSPFEFMDLWGLGAKIELS 450  
QY 417 VARGLEANGVDCDGETSLVVRGPDGKGLG---GGTVATHLDRHIAAMSLVWGLAEEKP 47  
Db 417 VARGLEANGVDCDGETSLVVRGPDGKGLG---GGTVATHLDRHIAAMSLVWGLAEEKP 47  
QY 447 VTVDATMIATSPFEFMDLWGLGAKIELS 447  
Db 447 VTVDATMIATSPFEFMDLWGLGAKIELS 447  
RESULT 12  
US-09-464-099a-70.ra  
; Sequence 5, Application us/09137440

Patent No. 6248876  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/137,440  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/833,485  
FILING DATE: 07-APR-1997  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-137-440-5

Query Match 83.1%; Score 1900.5; DB 4; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASSRATARKSGSLGTVIRIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHASPKPATARRSALGTEIRIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGGLLAPLDFGNATGCRLTMLGVGVDFDSTFI 120  
Db 61 RAMQAMGARIRKEDGTWIIDGVNGGGLLAPLDFGNATGCRLTMLGVGVDFDSTFI 120  
QY 121 GDASLKRPMGRVNLPLREMGVQVKSEDDGRLPVTLLRGPKTPTPIYRVPMSAQVKSAY 180  
Db 121 GDASLKRPMGRVNLPLREMGVQVKSEDDGRLPVTLLRGPKTPTPIYRVPMSAQVKSAY 180  
QY 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240  
Db 181 LLAGLNTPGTITVIEPTIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQVID 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTTGLILTLQEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTTGLILTLQEMGADIEVINPRLAGGED 300

```

Db 241 VPGDPSAPPLVAALLVEGSDVTIRNVLNPNRTGTLTLOEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNAGAAVATHLDRHRIAMSLVGLVSEN 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDGKGLGNAGAAVATHLDRHRIAMSLVGLVSEN 420
QY 421 VTVDATMIATSPFMDLMAGLCAKIELS 450
Db 418 VTVDSDNMIATSPFMDMMPGLGAKIELS 447

```

RESULT 13

```

US-09-137-440-7
; Sequence 7, Application US/09137440
; Patent No. 6248876
GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/137,440
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/833,485
; FILING DATE: 07-APR-1997
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-137-440-7

```

Query Match 83.1%; Score 1900.5; DB 4; Length 449;  
 Best Local Similarity 82.9%; Pred. No. 2.2e-167;

```

Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;
QY 1 MLHGASRPATARKSSGSLCTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db 1 MSHASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKKGDTWIIDGVNGGGLLAPAPLDFGNAAATGCRLTMLGLVGVYDFDSTFI 120
Db 61 RAMQAMGAKIRKKGDTWIIDGVNGGGLLAPAPLDFGNAAATGCRLTMLGLVGVYDMKTSFI 120
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSSEDGRLPVTLRGPKTPTTYRVPMASAOVKSAV 180
Db 121 GDASLSKRPMPGRVNLNPLREMGVQVEAADGDRMPLTLIGPKTANPTTYRVPMASAOVKSAV 180
QY 181 LLAGLNTPGITTVIEPIWTRDHTKMLQGFGANLTVETDADGVTRIIEGRGKLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIWTRDHTKMLQGFGANLTVETDADGVTRIIEGRGKLTGQVID 240
QY 241 VPGDPSAPPLVAALLVEGSDVTIRNVLNPNRTGTLTLOEMGADIEVLNARLAGGED 300
Db 241 VPGDPSAPPLVAALLVEGSDVTIRNVLNPNRTGTLTLOEMGADIEVLNARLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVYNGLEELRVKESDRLSA 360
QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNAGAAVATHLDRHRIAMSLVGLVSEN 420
Db 361 VARGLEANGVDCDEGETSLVVRGPDGKGLGNAGAAVATHLDRHRIAMSLVGLVSEN 420
QY 421 VTVDATMIATSPFMDLMAGLCAKIELS 450
Db 418 VTVDSDNMIATSPFMDMMPGLGAKIELS 447

```

RESULT 14

```

PCT-US91-06148A-5
; Sequence 5, Application PC/TUS9106148A
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06148A
; FILING DATE: 19910828
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/749,611
; FILING DATE: 28-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10535)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099

```

TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-06148A-5

Query Match 83.1%; Score 1900.5; DB 5; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASSRPATARKSGSLGTVRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60  
DB 1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGLLAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120  
DB 61 RAMQAMGARIRKEDGVWIIINGVNGCLLOPEALDFGNAAGTGARLTMLGLVGYTDMKTSFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQKSEGDRLPVTILRGPKTPTPIYRVPMSAQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQVEAADGRMPLTLIGPKTANPITYRVPMSAQVKSAY 180  
QY 181 LLAGLNTPGTITVIEPIVTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240  
DB 181 LLAGLNTPGTITVIEPIVTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300  
DB 241 VPGDPSSTAPFLVAALLVEGSDVTIRNVLNMPRTGLILTLQEMGADIEVLNARLAGGED 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAFAGATVWNGLEELRVKESDRLSA 360  
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAFAGATVWNGLEELRVKESDRLSA 360  
QY 361 VARGLEANGVDCCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRTAMSLVNGLAKEP 417  
DB 361 VARGLEANGVDCCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRTAMSLVNGLAKEP 417

RESULT 15  
T-US91-06148A-7  
Sequence 7, Application PC/TUS9106148A  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
TITLE OF INVENTION: Glycophosphate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/06148A  
FILING DATE: 19910828  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10535)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-06148A-7

Query Match 83.1%; Score 1900.5; DB 5; Length 449;  
Best Local Similarity 82.9%; Pred. No. 2.2e-167;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;  
QY 1 MLHGASSRPATARKSGSLGTVRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60  
DB 1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFPMFGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIIDGVNGGLLAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120  
DB 61 RAMQAMGARIRKEDGVWIIINGVNGCLLOPEALDFGNAAGTGARLTMLGLVGYTDMKTSFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQKSEGDRLPVTILRGPKTPTPIYRVPMSAQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQVEAADGRMPLTLIGPKTANPITYRVPMSAQVKSAY 180  
QY 181 LLAGLNTPGTITVIEPIVTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240  
DB 181 LLAGLNTPGTITVIEPIVTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGOVID 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300  
DB 241 VPGDPSSTAPFLVAALLVEGSDVTIRNVLNMPRTGLILTLQEMGADIEVLNARLAGGED 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAFAGATVWNGLEELRVKESDRLSA 360  
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAFAFAGATVWNGLEELRVKESDRLSA 360  
QY 361 VARGLEANGVDCCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRTAMSLVNGLAKEP 417  
DB 361 VARGLEANGVDCCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRTAMSLVNGLAKEP 417

Search completed: August 15, 2002, 13:59:17  
Job time: 91 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:46 ; Search time 22.39 Seconds  
(without alignments)  
496.366 Million cell updates/sec

Title: US-09-464-099A-3

Perfect score: 2288

Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLWGLGKLTSLDTRAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Local number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	1	US-08-476-008-3
2	2288	100.0	455	1	US-08-306-063-3
3	2288	100.0	455	1	US-08-833-485-3
4	2288	100.0	455	4	US-09-137-440-3
5	2288	100.0	455	5	PCT-US91-06148A-3
6	1906.5	83.3	449	1	US-08-476-008-5
7	1906.5	83.3	449	1	US-08-476-008-7
8	1906.5	83.3	449	1	US-08-306-063-5
9	1906.5	83.3	449	1	US-08-306-063-7
10	1906.5	83.3	449	1	US-08-833-485-5
11	1906.5	83.3	449	1	US-08-833-485-7
12	1906.5	83.3	449	4	US-09-137-440-5
13	1906.5	83.3	449	4	US-09-137-440-7
14	1906.5	83.3	449	5	PCT-US91-06148A-5
15	1906.5	83.3	449	5	PCT-US91-06148A-7
16	867.5	37.9	447	1	US-08-476-008-67
17	867.5	37.9	447	1	US-08-306-063-67
18	867.5	37.9	447	1	US-08-833-485-67
19	867.5	37.9	447	4	US-09-137-440-67
20	806	35.2	443	1	US-08-476-008-69
21	806	35.2	443	1	US-08-306-063-69
22	806	35.2	443	1	US-08-833-485-69
23	806	35.2	443	4	US-09-137-440-69
24	803.5	35.1	427	4	US-09-325-881-2
25	798.5	34.9	427	2	US-08-896-345-2
26	798.5	34.9	427	4	US-09-226-091-2
27	785	34.3	428	1	US-08-476-008-42

28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appl
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appl
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appl
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appl
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

#### ALIGNMENTS

RESULT 1  
US-08-476-008-3  
; Sequence 3, Application US/08476008  
; Patent No. 5627061

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Padgett, Stephen R.

; APPLICANT: Stallings, William C.

; TITLE OF INVENTION: Glyphosate Tolerant

; NUMBER OF SEQUENCES: 69

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F

; STREET: 700 Chesterfield Village Parkway

; CITY: St. Louis

; STATE: Missouri

; COUNTRY: USA

; ZIP: 63198

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,008

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/306,063

; FILING DATE: 13-SEP-1994

; APPLICATION NUMBER: US 07/749,611

; FILING DATE: 28-AUG-1991

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/576,537

; FILING DATE: 31-AUG-1990

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Hoerner Jr., Dennis R.

; REGISTRATION NUMBER: 30,914

; REFERENCE/DOCKET NUMBER: 38-21(10660)A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (314)537-6099

; TELEFAX: (314)537-6047

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 455 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-476-008-3

Query Match 100.0%; Score 2288; DB 1; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.6e-203; Indels 0; Gaps 0;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPARATARKSGSLGTVRIPGDKSISHRSPMFGGLASGETRITGLLEGEVDINTG 60  
DB 1 MSHGASSRPARATARKSGSLGTVRIPGDKSISHRSPMFGGLASGETRITGLLEGEVDINTG 60

QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120  
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120

QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPPTITRYRVPMSAQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPPTITRYRVPMSAQVKSAY 180

QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVTRTLRGRGKLTQGVID 240  
DB 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVTRTLRGRGKLTQGVID 240

QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300  
DB 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300

QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPTILAVAAAFAGCATVMNGLEELRVKESDRLSA 360  
DB 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPTILAVAAAFAGCATVMNGLEELRVKESDRLSA 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSLVLMGLVSENP 420  
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSLVLMGLVSENP 420

QY 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455  
DB 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455

RESULT 2  
US-08-306-063-3  
Sequence 3, Application US/08306063  
Patent No. 5633435  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/306,063  
APPLICATION NUMBER: US/08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-306-063-3

Query Match 100.0%; Score 2288; DB 1; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.6e-203; Indels 0; Gaps 0;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPARATARKSGSLGTVRIPGDKSISHRSPMFGGLASGETRITGLLEGEVDINTG 60  
DB 1 MSHGASSRPARATARKSGSLGTVRIPGDKSISHRSPMFGGLASGETRITGLLEGEVDINTG 60

QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120  
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLPEAPLDFGNAATGCRLTMLGLVGVYDFDSTFI 120

QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPPTITRYRVPMSAQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPPTITRYRVPMSAQVKSAY 180

QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVTRTLRGRGKLTQGVID 240  
DB 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVTRTLRGRGKLTQGVID 240

QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300  
DB 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGLILTLQEMGADIEVINPRLAGGED 300

QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPTILAVAAAFAGCATVMNGLEELRVKESDRLSA 360  
DB 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPTILAVAAAFAGCATVMNGLEELRVKESDRLSA 360

QY 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSLVLMGLVSENP 420  
DB 361 VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSLVLMGLVSENP 420

QY 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455  
DB 421 VTVDATMIATSPFPEMDLMAGLAKIELSDTKAA 455

RESULT 3  
US-08-833-485-3  
Sequence 3, Application US/08833485  
Patent No. 5804425  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis

STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-833-485-3

Query Match 100.0%; Score 2288; DB 1; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.6e-203;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
DB 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDPGNAATGRLTMGLVGVYDFDSTFI 120  
DB 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDPGNAATGRLTMGLVGVYDFDSTFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTILRGPKTPTPIYRVPMAOQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTILRGPKTPTPIYRVPMAOQVKSAY 180  
QY 181 LLAGLNTPGITTVIEPIWTRDHTERKMLQFGANLTVETDADGVRTIRLEGRKLTQVID 240  
DB 181 LLAGLNTPGITTVIEPIWTRDHTERKMLQFGANLTVETDADGVRTIRLEGRKLTQVID 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNPNTRGLIILTLOEMGADIEVINPRLAGGED 300  
DB 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNPNTRGLIILTLOEMGADIEVINPRLAGGED 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
QY 361 VANGKLNGVDCDEGETSLVWRGRPDGKGLGNASGAATVHLDHRTAMSFVNGLYSENP 420  
DB 361 VANGKLNGVDCDEGETSLVWRGRPDGKGLGNASGAATVHLDHRTAMSFVNGLYSENP 420  
QY 421 VTVDATMTATSFPEFMDLMAGLGAKIELSDTKAA 455

DB 421 VTVDATMTATSFPEFMDLMAGLGAKIELSDTKAA 455  
RESULT 4  
US-09-137-440-3  
Sequence 3, Application US/09137440  
Patent No. 6248876  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyophosphate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/137,440  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/833,485  
FILING DATE: 07-APR-1997  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 455 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-137-440-3

Query Match 100.0%; Score 2288; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.6e-203;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
DB 1 MSHGASSRATARKSSGLSGTVRIPGDKSISHRSMFGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDPGNAATGRLTMGLVGVYDFDSTFI 120  
DB 61 KAMQAMGARIRKEGDWTIIDGVNGGGLLAPEAPLDPGNAATGRLTMGLVGVYDFDSTFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTILRGPKTPTPIYRVPMAOQVKSAY 180  
DB 121 GDASLTKRPMGRVNLPLREMGVQVKSSEDGRLPVTILRGPKTPTPIYRVPMAOQVKSAY 180

Query Match	100.0%	Score 2288;	DB 5;	Length 455;
Best Local Similarity	100.0%	Pred. No. 2.6e-203;		

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Hoerner Jr., Dennis R.
: REGISTRATION NUMBER: 30,914
: REFERENCE/DOCKET NUMBER: 38-21(10660)A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314)537-6099
: TELEFAX: (314)537-6047
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 449 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-476-008-5

```

Query Match	83.3%	Score 1906.5	DB 1	Length 449
Best Local Similarity	83.1%	Pred. No. 4.7e-168		
Matches 374	Conservative 32	Mismatches 41	Indels 3	Gaps
1	MSHGASSRPATARKSGLSGTVIRPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTG	60		
Db	:       :     :       :       :       :       :       :       :			
1	MSHSAPKPATARSEALTEGRIIPGDKSISHRSPFPGGLASGETRITGLLEGEDVINTG	60		
Db	:       :       :       :       :       :       :       :       :			
61	RAMQAMGARTRKEGDYWIIDGVNGGGLLAPAPLDFGNAAATGRLPMGLVGVYDFDSTFI	120		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
61	RAMQAMGAKTRKEGDYWIINGVNGGGLLOPEAALDFGNAGTAGTARLPMGLVGVYDMKTSFI	120		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
121	GDASLTKRPMGRVNLNPREMGVOVKSEDDGRLPVTILRGPKTPPIPIYRVPMAQAQVKSAV	180		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
121	GDASLTKRPMGRVNLNPREMGVOVEAADGRMPLTLLGPKTANPIYRVPMAQAQVKSAV	180		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
181	LLAGLNTPGTIVTIEIPIMTRDHTHEKMLQFGANLTVETDADGVRTIRLEGRCGLTKGQVID	240		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
181	LLAGLNTPGTIVTIEIPVMTDRDHTHEKMLQFGADLTVDKDGVRHRIITCGQKLVGQTTID	240		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
241	VPDPSSTAPFLVAALLVPGSDVTILNVLNPNPTGTGLLTLOEMGADIETVNPRLAGGED	300		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
241	VPDPSSTAPFLVAALLVPGSDVTIRNVLNPNPTGTGLLTLOEMGADIEVLNARLAGGED	300		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
301	VADLRVRSSTLKGVTVPEDRAPSMIDEXPI LAVAAAFAGATVMNGLEELRVKESDRLSA	360		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
301	VADLRVRSASLKGWVYPPERAPSMIDEXPIVLAATAAFAGETVMNGLDELRVKESDRLSA	360		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
361	VANGLKLVGDCDEGETSLVVRGRPGDKGLGNASGAAVATHLDHRTAMSLFVYMGVLENP	420		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
361	VARGLEANGVDCYTEGEMSLTVRGRPGDKGLG---GGTVATHLDHRTAMSLFVYMGVLENP	417		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
421	VTVDADPMIATSPPEFMDLMAGLCAKIELS	450		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			
418	VTVDSDNNIATSPPEFMDMMPLGAKIELS	447		
QY	:       :       :       :       :       :       :       :       :			
Db	:       :       :       :       :       :       :       :       :			

```

RESULT      7
US-08-476-008-7
Sequence 7, Application US/08476008
Patent No. 5627061
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate
TITLE OF INVENTION: 5-Enolpyruv
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner,
STREET: 700 Chesterfield Vill
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,008
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-008-7

```

Query Match	83.3%;	Score 1906.5;	DB 1;	Length 449;
Best Local Similarity	83.1%;	Pred. No. 4.7e-168;		
Matches 374;	Conservative 32;	Mismatches 41;	Indels 3;	Gaps 1;
Qy	1	MSHGASSRPATARKSSGLSGTVRIPGOKSISHRFSMFQGLASGETRITGLLEGEDVINTG	60	
Db	1	MSHSASPKPATARSEALTGEIRIPGDKSISHRFSMFQGLASGETRITGLLEGEDVINTG	60	
Qy	61	KAMQAMGARIRKEGDTWIIDGVGNGGLLAPAPLDFGNAATGCRITMGLVGVYDFDSTFI	120	
Db	61	KAMQAMGAKIRKEGDVWIIINGVNGCLLPQEAALDFGNAGTGARITMGLVGTYDMKTSFI	120	
Qy	121	GDASITRKPMGRVLNPLREMGVQVKSQEDRLPVTLRGPKTPTPTTVRPNMAAQVKSAV	180	
Db	121	GDASISRKPMGRVLNPLREMGVQVEADGRMLTLTGPKTANPTTVRPNMAAQVKSAV	180	
Qy	181	LLAGLNTPGTIVTTEPIMTRDHTKMKLGQGANITVETDADGVRTIRLEGRGKLTQGVID	240	
Db	181	LLAGLNTPGVTIVTTEPVMVTRDHTKMKLGFGADITVETDKDGVHRITGTGGKLGQVID	240	
Qy	241	VPDPSSTAFPLVAALLVPGSDVITILNVMNPTRTGLTLTLOEMGADIEVINPLAGGED	300	
Db	241	VPDPSSTAFPLVAALLVPGSDVITIRNVMNPTRTGLTLTLOEMGADIEVLNARLAGED	300	
Qy	301	VADLRVRSSTLKVTVVPEDRAPSMIDEVPIIAVAAAFAGATVWNGLEBELRVKESDRUSA	360	
Db	301	VADLRVRASKLKVVPVPERAPSMIDEVPIVIAIAASFAGETVMDGLDELRVKESDRUAA	360	
Qy	361	VANGKLKNGVDCDEGETSLVVYRGPRDGGKLGNAASAAVATHLDHRIAMSFVLMGLVSEN	420	
Db	361	VARGLEANGVDCETEGENSLTVYRGPRDGGKLG---GGTVATHLDHRIAMSFVLMGLAAEKP	417	
Qy	421	VTVDDATMIATSFPEFMDLMAGLAKIELS	450	
Db	418	VTVDDSNMIATSFPEFMDMPGLGAKIELS	447	

RESULT 8  
US-08-306-063-5

```

; Sequence 5, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-5

```

```

Query Match      83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRATARKSGSLGTVRIPDKSISHRSFMFGGLASGTRITGLLEGEVDINTG 60
DB 1 MSHSASPKPATARRSEALTGEIRIPDKSISHRSFMFGGLASGTRITGLLEGEVDINTG 60

QY 61 KANQMGARTRKGGDWIDGVNGGLLAPDFGNATGCRMTGLVGVYDFDSTFI 120
DB 61 KANQMGARTRKGGDWIDGVNGGLLAPDFGNATGCRMTGLVGVYDFDSTFI 120

QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180
DB 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPTIYRVPMSAQVKSAY 180

QY 181 LLAGLNTPGTIVTIEPTWTRDHEKMLQCGANLTVETDADGVTRTLRGRGKLGTQVID 240
DB 181 LLAGLNTPGTIVTIEPTWTRDHEKMLQCGANLTVETDADGVTRTLRGRGKLGTQVID 240

QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNRTGLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNRTGLILTLQEMGADIEVINPRLAGGED 300

```

```

QY 301 VADLRVRSSTIKGVTVPPEDRAPSMIDEPYILAVAAAFAGATVWNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTIKGVTVPPEDRAPSMIDEPYILAVAAAFAGATVWNGLEELRVKESDRLSA 360

QY 361 VANGKLINGVDCDEGETSLVVGRPDGKGLGNASCAAVATHLDHRIAMSLVWGLVSENP 420
DB 361 VANGKLINGVDCDEGETSLVVGRPDGKGLGNASCAAVATHLDHRIAMSLVWGLVSENP 420

QY 421 VTVDATMIATSPFPEMDLMAGLCAKIELS 450
DB 421 VTVDATMIATSPFPEMDLMAGLCAKIELS 450

QY 450 VTVDATMIATSPFPEMDLMAGLCAKIELS 450
DB 450 VTVDATMIATSPFPEMDLMAGLCAKIELS 450

```

```

RESULT 9
US-08-306-063-7
; Sequence 7, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-7

```

```

Query Match      83.3%; Score 1906.5; DB 1; Length 449;
Best Local Similarity 83.1%; Pred. No. 4.7e-168;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRATARKSGSLGTVRIPDKSISHRSFMFGGLASGTRITGLLEGEVDINTG 60
DB 1 MSHSASPKPATARRSEALTGEIRIPDKSISHRSFMFGGLASGTRITGLLEGEVDINTG 60

```

[illegible]

RESULT 10  
US-08-833-485-5  
Sequence 5, Application US/08833485  
Patent No. 5804425  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgette, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914

```

; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-485-5

Query Match      83.3%; Score 1906.5, 5.
Best Local Similarity 83.1%; Pred. No. 4.
Matches 374; Conservative 32; Mismatches

QY      1  MSHGASSRRPATARKSGSLGCTVTRIPGDKSISHRFSFMM
      ||| ||| :|||||: ||| : |||||:|||||:|||||:
DB      1  MSHSASPKPATARRSALGTGEIRIPGDKSISHRFSFMM
      ||| ||| :|||||: ||| : |||||:|||||:|||||:
QY     61  KAMQAMGARIRKEGDTWIIDGVNGGGLLAPPEAPLDVF
      |||||:|||||:|||||:|||||:|||||:|||||:
DB     61  RAMQAMGAKIRKEGDYIIINGVNGCQLQPEAAADVF
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    121  GASLTKRPMGRVNLPLREMGVQVKSEDDRLPVTLL
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    121  GASLSKRPMPGRVNLPLREMGVQVEAADGDRMPLTL
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    181  LLAGLNTPGTITTVIEPTMDTRDTEKMLQGFGANLTVTI
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    181  LLAGLNTPGTITTVIEPWTMDTRDTEKMLQGFGANLTVTI
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    241  VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPRTGT
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    241  VPGDPSSTAPFLVAALLVEGSDVTIRNVLMNPRTGT
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    301  VADLRVRSSTLKGTVTPEDRAPSMIDEPYILAVAAA
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    301  VADLRVRASKLGTVVPPERAPSMIDEPYILAVAAAS
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    361  VANGKLINGVDCDEGETSLVVRGPDGKGLGNASGA
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    361  VARGLEANGVDCTEGEMSLTVRGPDKGLG---GG
      |||||:|||||:|||||:|||||:|||||:|||||:
QY    421  VTVDATMIATSPPEFMDLMAGLAKIELS 450
      |||||:|||||:|||||:|||||:|||||:|||||:
DB    418  VTVDSSNMIAITSPPEFMDMPGLGAKIELS 447
      |||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 11
US-08-833-485-7
; Sequence 7, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enoipyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-833-485-7

Query Match 83.3%; Score 1906.5; DB 1; Length 449;  
Best Local Similarity 83.1%; Pred. No. 4.7e-168;  
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRATARKSGSLCTVRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHSASPKPATARRSEALTEGIRPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIDGVNGGLLAPPEALDFGNATGCRLTMLGVGVYDFDSTFI 120  
Db 61 RAMQAMGAKIRKEDGVWIIINGVNGCLLQPEALDFGNAGTGARLTMLGVGVYDMKTSFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPIYRVPMASQAQVKSAY 180  
Db 121 GDASLSKRPMGRVNLPLREMGVQVEAADGRMPLTLIGPKTANPIYRVPMASQAQVKSAY 180  
181 LLAGLNPPTGTTTVEIPTMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGOVID 240  
181 LLAGLNPPTGTTTVEIPVMTDRHTKMLQGFADLTVEDTKDGVRRHIRTGQGLVGGTID 240  
QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNVLNMPRTTGLILTLOEMGADIEVINPRLAGGED 300  
Db 241 VPGDPSSTAPPLVAALLVPGSDVTIRNVLNMPRTTGLILTLOEMGADIEVINARLAGGED 300  
QY 301 VADLRVRSSTLKGVTPEDRAPSMIDEPYTLAVAAFAEAGTVMNGLEELRVKESDRLSA 360  
Db 301 VADLRVRASKLGVVPPERAPSMIDEPYVLAIAAFAEGETVMNGLDLRLVKESDRLAA 360  
QY 361 VANGKLUNGVDCEGETSLVVRGPDGKGLGNASGAARVATHLDRHTAMSLVNGLVSENP 420  
Db 361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDRHTAMSLVNGLVAAEKP 417  
QY 421 VTVDATMIATSPPEFMDLMAGLAKIELS 450  
Db 418 VTVDSSNNIATSPPEFMDMPGLGAKIELS 447

RESULT 12  
US-09-137-440-5  
; Sequence 5, Application.US/09137440

Patent No. 6248876  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B&F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/137,440  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/833,485  
FILING DATE: 07-APR-1997  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-137-440-5

Query Match 83.3%; Score 1906.5; DB 4; Length 449;  
Best Local Similarity 83.1%; Pred. No. 4.7e-168;  
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRATARKSGSLCTVRIPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHSASPKPATARRSEALTEGIRPGDKSISHRSFMEFGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMGARIRKEDGTWIDGVNGGLLAPPEALDFGNATGCRLTMLGVGVYDFDSTFI 120  
Db 61 RAMQAMGAKIRKEDGVWIIINGVNGCLLQPEALDFGNAGTGARLTMLGVGVYDMKTSFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQKSDGDRPLVTLRGPKTPPIYRVPMASQAQVKSAY 180  
Db 121 GDASLSKRPMGRVNLPLREMGVQVEAADGRMPLTLIGPKTANPIYRVPMASQAQVKSAY 180  
181 LLAGLNPPTGTTTVEIPTMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGOVID 240  
181 LLAGLNPPTGTTTVEIPVMTDRHTKMLQGFADLTVEDTKDGVRRHIRTGQGLVGGTID 240  
QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNVLNMPRTTGLILTLOEMGADIEVINPRLAGGED 300



Db 241 VPGDPSSTAPLVAALLVSGSDVTIRNVLNPNRTGLTLTQEMGADIEVNLNARLAGGED 300  
QY 301 VADLRVSTLKGVTVPEDRAPSMDIDEYPTLAVAAFAEGATYMGLEELRVKESDRLSA 360  
Db 301 VADLRVRASKLKGVTVPEDRAPSMDIDEYPTLAVAAFAEGATYMGLEELRVKESDRLSA 360  
QY 361 VANGKLVGDCDEGETSLVVRGPRDCKGLGNASGAATVATHLDRHTAMSLVGLVSEN 420  
Db 361 VAGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDRHTAMSLVGLVSEN 417  
QY 421 VTVDATMIATSPFPEMDLMAGLGAKIELS 450  
Db 418 VTVDSDNMIATSPFPEMDMMPGLGAKIELS 447

RESULT 13  
US-09-137-440-7  
; Sequence 7, Application US/09137440  
Patent No. 6248876  
GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
; STREET: 700 Chesterfield Village Parkway  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/09/137,440  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/833,485  
; FILING DATE: 07-APR-1997  
; APPLICATION NUMBER: US 08/306,063  
; FILING DATE: 13-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(15117)A  
; TELEPHONE: (314)737-6099  
; TELEFAX: (314)737-6047  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 449 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-137-440-7

Query Match 83.3%; Score 1906.5; DB 4; Length 449;  
Best Local Similarity 83.1%; Pred. No. 4.7e-168;

Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;  
QY 1 MSHGASSRPATARKSGSLGTVIRPGDKSIHSRPFMGGLASGETRITGLLEGEDVINTG 60  
Db 1 MSHSASPKPATARRSEALTGEIRIPGDKSIHSRPFMGGLASGETRITGLLEGEDVINTG 60  
QY 61 KAMQAMCARIRKESGDWIIIDGVNGGGLLAPLDFCNATGCLRTMGLVGVYDFDSTFI 120  
Db 61 RAMQAMGAKIRKESGDWIIIDGVNGGGLLQPEAALDFGNAGTGARLTMLGLVGTYDMKTSFI 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQVKSQDGLRPLVTLRGPKTPTPTIYRVPMASAOQVKS 180  
Db 121 GDASLTKRPMGRVNLPLREMGVQVKSQDGLRPLVTLRGPKTPTPTIYRVPMASAOQVKS 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVTRTLRBOGKLTQGV 240  
Db 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFADLTETDKDGVHRIITQGGKLVGTQ 240  
QY 241 VPGDPSSTAPLVAALLVSGSDVTIRNVLNPNRTGLTLTQEMGADIEVNLNARLAGGED 300  
Db 241 VPGDPSSTAPLVAALLVSGSDVTIRNVLNPNRTGLTLTQEMGADIEVNLNARLAGGED 300  
QY 301 VADLRVSTLKGVTVPEDRAPSMDIDEYPTLAVAAFAEGATYMGLEELRVKESDRLSA 360  
Db 301 VADLRVRASKLKGVTVPEDRAPSMDIDEYPTLAVAAFAEGATYMGLEELRVKESDRLSA 360  
QY 361 VANGKLVGDCDEGETSLVVRGPRDCKGLGNASGAATVATHLDRHTAMSLVGLVSEN 420  
Db 361 VAGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDRHTAMSLVGLVSEN 417  
QY 421 VTVDATMIATSPFPEMDLMAGLGAKIELS 450  
Db 418 VTVDSDNMIATSPFPEMDMMPGLGAKIELS 447

RESULT 14  
PCT-US91-06148A-5  
; Sequence 5, Application PC/TUS9106148A  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
; STREET: 700 Chesterfield Village Parkway  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/06148A  
; FILING DATE: 19910828  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(10535)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-6099

TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-06148A-5

Query Match 83.3%; Score 1906.5; DB 5; Length 449;  
Best Local Similarity 83.1%; Pred. No. 4.7e-168;  
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

1 MSHGASSRPATARKSGSLGTVIRPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
1 MSHSASPATARRSEALTGETIRPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCRLTMLGVGYDFDSTFI 120  
61 RAMQAMGAKIRKEGDTWIIDGVNGGLLAPLDFGNAATGCRLTMLGVGYDFDSTFI 120  
121 GSAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180  
121 GDSLSKRPMPGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180  
181 LLAGLNTPGTIVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGOVID 240  
181 LLAGLNTPGTIVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGOVID 240  
241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300  
241 VPGDPSSTAFPLVAALLVPGSDVTIRNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300  
301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360  
301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLHRIAMSLVGLVSENP 420  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLHRIAMSLVGLVSENP 420  
421 VTDDATMIATSPFPMDLMLAGLAKIELS 450  
418 VTDDSNMIATSPFPMDLMLAGLAKIELS 447

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/749,611  
FILING DATE: 28-AUG-1991  
PRIOR APPLICATION DATA: US 07/576,537  
APPLICATION NUMBER: 31-AUG-1990  
FILING DATE: 31-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10535)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-06148A-7

TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-06148A-5

Query Match 83.3%; Score 1906.5; DB 5; Length 449;  
Best Local Similarity 83.1%; Pred. No. 4.7e-168;  
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

1 MSHGASSRPATARKSGSLGTVIRPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
1 MSHSASPATARRSEALTGETIRPGDKSISHRSPFMFGGLASGETRITGLLEGEDVINTG 60  
61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCRLTMLGVGYDFDSTFI 120  
61 RAMQAMGAKIRKEGDTWIIDGVNGGLLAPLDFGNAATGCRLTMLGVGYDFDSTFI 120  
121 GSAISLTKRPMGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180  
121 GDSLSKRPMPGRVNLPLREMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMSAQVKSAY 180  
181 LLAGLNTPGTIVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGOVID 240  
181 LLAGLNTPGTIVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGRGKLTGOVID 240  
241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300  
241 VPGDPSSTAFPLVAALLVPGSDVTIRNVLNPNTRTGLILTQEMGADIEVINPRLAGGED 300  
301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360  
301 VADLRVRSSTLKGTVVPEDRAPSMIDEPILAVAAFAFAGATVMNGLEELRVKESDRLSA 360  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLHRIAMSLVGLVSENP 420  
361 VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAATAVATHLHRIAMSLVGLVSENP 420  
421 VTDDATMIATSPFPMDLMLAGLAKIELS 450  
418 VTDDSNMIATSPFPMDLMLAGLAKIELS 447

RESULT 15  
PCT-US91-06148A-7  
Sequence 7, Application PC/TUS9106148A  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
TITLE OF INVENTION: Glycosylated Tolerant  
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/06148A  
FILING DATE: 19910828  
CLASSIFICATION: 800

Search completed: August 15, 2002, 13:59:16  
Job time: 90 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: August 15, 2002, 14:05:43 ; Search time 51.67 Seconds  
(without alignments)  
1523.373 Million cell updates/sec

Title: US-09-464-099A-70  
Perfect score: 2288  
Sequence: 1 MLHGASSRATARKSSGLSG.....FMDLMAGLGNKIELSDTKAA 455

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

11 number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-virus:\*  
13: SP-vertebrate:\*  
14: SP-unclassified:\*  
15: SP-rvirus:\*  
16: SP-bacteriap:\*  
17: SP-archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2098	91.7	455	16 Q92SV5	Q92SV5 rhizobium m
2	1855.5	81.1	480	2 Q9AGV2	Q9AGV2 brucella ab
3	1783	77.9	452	16 Q98CC1	Q98CC1 rhizobium l
4	1135	49.6	443	16 Q9A2H2	Q9A2H2 caulobacter
5	989.5	43.2	746	16 Q9H269	Q9H269 pseudomonas
6	912.5	39.9	454	16 Q9PB21	Q9PB21 xylella fas
7	884	38.6	431	16 Q9KCA6	Q9KCA6 bacillus ha
8	820	35.8	430	16 Q99Z83	Q99Z83 streptococ
9	815.5	35.6	428	16 Q92A85	Q92A85 listeria in
10	781.5	34.2	428	2 Q9ANY6	Q9ANY6 enterococu
11	610	26.7	432	16 Q99U25	Q99U25 staphylococ
12	470.5	20.6	207	2 Q9RH28	Q9RH28 pseudomonas
13	470.5	20.6	439	17 Q9HQC1	Q9HQC1 halobacteri
14	416.5	18.2	440	2 Q9L213	Q9L213 streptomyce
15	401.5	17.5	426	16 Q9KR80	Q9KR80 vibrio chol
16	379	16.6	427	2 Q93ED4	Q93ED4 yersinia ru

17	351.5	15.4	428	16 Q97KM2	Q97KM2 clostridium
18	324.5	14.5	438	2 Q9K4A7	Q9K4A7 streptomyce
19	324.5	14.2	447	16 Q9K9D5	Q9K9D5 bacillus ha
20	306	13.4	516	10 Q946V0	Q946V0 dicliptera
21	302	13.2	521	10 Q9FVP6	Q9FVP6 arabidopsis
22	299	13.1	444	10 Q24566	Q24566 zea mays (m
23	298	13.0	444	10 Q24566	Q24566 dicliptera
24	297.5	13.0	511	10 Q93VK6	Q93VK6 oryza sativ
25	293.5	12.8	445	8 Q95AK0	Q95AK0 eleusine in
26	292.5	12.8	445	8 Q95AK1	Q95AK1 eleusine in
27	285	12.5	433	16 Q9JTT3	Q9JTT3 neisseria m
28	280	12.2	414	17 Q980I5	Q980I5 sulfolobus
29	275	12.0	433	16 Q9JYU1	Q9JYU1 neisseria m
30	261	11.4	408	17 Q96Y91	Q96Y91 sulfolobus
31	260.5	11.4	427	17 Q9YC47	Q9YC47 aeropyrum p
32	229	10.0	332	2 Q9ZEQ0	Q9ZEQ0 actinobacil
33	223	9.7	410	17 Q978S3	Q978S3 thermoplasm
34	216.5	9.5	391	10 Q80428	Q80428 oryza sativ
35	206.5	9.0	448	2 Q9L1U5	Q9L1U5 streptomyce
36	203.5	8.9	446	2 Q9S0N1	Q9S0N1 streptomyce
37	185.5	8.1	419	16 Q99Z78	Q99Z78 streptococ
38	184.5	8.1	417	16 Q9K1Q9	Q9K1Q9 neisseria m
39	184	8.0	418	16 Q97DD9	Q97DD9 clostridium
40	183.5	8.0	417	16 Q9JWS7	Q9JWS7 neisseria m
41	182	8.0	428	16 Q9K6E5	Q9K6E5 bacillus ha
42	181	7.9	419	2 Q9S5W5	Q9S5W5 escherichia
43	179	7.8	347	10 Q9AT37	Q9AT37 lolium rigi
44	179	7.8	423	16 Q927U1	Q927U1 listeria in
45	178.5	7.8	417	2 Q9EXE3	Q9EXE3 mycobacteri

#### ALIGNMENTS

RESULT 1

Q92SV5 PRELIMINARY; PRT; 455 AA.  
AC Q92SV5;  
DT 01-DEC-2001 (TREMREL. 19, Created)  
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)  
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)  
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN (EC 2.5.1.19).  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21368234; PubMed=11474104;  
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G., Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M., Vorhoeiter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";  
RL Science 293:668-672(2001).  
DR EMBL; AL591783; CAC41690.1; -.  
KW Transferrase; Complete proteome.  
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 91.7%; Score 2098; DB 16; Length 455;

Best Local Similarity 90.5%; Pred. No. 2.3e-119;

Matches 412; Conservative 18; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MLHGASSRATARKSSGLSGTIRIPDGSISHRSPMFGGLASGETRITGLLGEDVINTG 60

Db	1	MSHGSNPRPATAKSSDLKGTURIPGDKSISHRSFMEGGLAAGETRIITGLLEGEDVIINTG	60
Qy	61	KAMQAMGARIRKEGDTWIIDGVNGGALLAPEAPLDFGNAAATCRLTMTGLVGYVDFDSTFI	120
Db	61	KAMQAMGARIRKEGDTWIIDGVNGGALLAPEAPLDFGNAAATCRLTMTGLVGYVDFDSTFI	120
Qy	121	GDASLTKRPMGRVNLNPREMGVOVKSEDDRLPVTILRGPKTPTITYRVPMSAAQVKSAV	180
Db	121	GDASLTKRPMGRVLDPLREMGVOVKSAGEDRLPVTILRGPKTNPITYRVPMSAAQVKSAV	180
Qy	181	LLAGLNTPGTITTVIEPIMTDRDHTKMLQFGGANLTVETDADGVPIRLEGRKLTGQVID	240
Db	181	LLAGLNTPGTITTVIEPIMTDRDHTKMLQFGGANLTVETDAEGVYRTIRLEGRKLTGQVID	240
Qy	241	VPGDPSSTAPLVAALLVPGSDVTIILNVLNPNRTGLILTLOEMGADIEVINPRLAGGED	300
Db	241	VPGDPSSTAPLVAAGLVPGSDVTIILNVLNPNRTGLILTLOEMGANIEVMNKRLAGGED	300
Qy	301	VADLVRSSSTLKGVTVPEDRAPSMIDEXPILAFAAAFAEGATVMNGLEELRVKESDRLSA	360
Db	301	VADLVRHSELKGVTVPEDRAPSMIDEXPVLAFAAFAEGTVMNGLEELRVKESDRLSA	360
Qy	361	VANGLKNGVDCDEGETSLVYVRPGDKGLGNASGAATAHLDHRTAMSFVLVNGLVSENP	420
Db	361	VADGLKNGVDCDEGEASLVYVRPGGKGLKISGGQVKTHLDRHTAMSFVLVNGLASEHP	420
Qy	421	VTVDATMIATSPFEFMDLMAGLCAKTELSDTRAA	455
Db	421	VTVDATMIATSPFEFMDLMAGLCAKTEEAENKAA	455

RESULT	2	
Q9AGV2		
ID	Q9AGV2	PRELIMINARY; PRT: 480 AA.
AC	Q9AGV2;	
DT	01-JUN-2001	(TREMblrel. 17, Created)
DT	01-JUN-2001	(TREMblrel. 17, Last sequence update)
DT	01-OCT-2001	(TREMblrel. 18, Last annotation update)
DE	5-ENOLPYRUVYL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).	
GN	AROAN.	

Qy	121	GDASLTKRPMGRVNLNPLREMGVOVKSESDGDRULPVTLLRGGPKTPTPTIYRVPMASAOQKSAV	180
Db	151	GDASLTKRPMGRVNLNPLREMGVOVKSESDGDRULPVTLLRGGPKTPTPTIYRVPMASAOQKSAV	210
Qy	181	LLAGLNTPGTITVTIEPTMTDRHTEKMLQGFAGNLTVETDADGVRTIRLEGRKLTGQVID	240
Db	211	LLAGLNTPGTITVTIEPTMTDRHTEKMLQGFAGNLTVETDADGVRTIRLEGRKLTGQVID	270
Qy	241	VPGDSPSTAFPLVAALLVPGSDVTIILNVLNPNPRTTGLIITLQEMGADIIEVINPRLAGGED	300
Db	271	VPGDSPSTAFPLVAALLVPGSDVTIILNVLNPNPRTTGLIITLQEMGADIIEVINPRLAGGED	330
Qy	301	VADLRVRSSTLKGVTVPEDRAPSMIDEPYPLTAVAAAFAGCATVMNGLEELRVKESDRLSA	360
Db	331	VADLRVRSSTLKGVTVPEDRAPSMIDEPYPLTAVAAAFAGCATVMNGLEELRVKESDRLSA	390
Qy	361	VANGLKUNGVDCDEGETSLTVYVRGPDKGLGNAGSAAVATHLDRHTAMFSLVNGVISENP	420
Db	391	VARGLEANGVDCDEGETSLTVYVRGPDKGLGNAGSAAVATHLDRHTAMFSLVNGVISENP	447
Qy	421	VTVDATMIATSPFPEFMDLNAAGLAKI	447
Db	448	VTVDSTMIATSPFPEFMDLNAAGLAKI	474
RESULT	3		
Q98CC1		PRELIMINARY;	PRT;
ID	Q98CC1		452 AA.
AC	Q98CC1;		
DT	01-OCT-2001 (TReMBLrel. 18, Created)		
DT	01-OCT-2001 (TReMBLrel. 18, Last sequence update)		
DT	01-OCT-2001 (TReMBLrel. 18, Last annotation update)		
DF	3-PROSPOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.		

RESULT	3
Q98CC1	
ID	PRELIMINARY; PRT; 452 AA.
AC	Q98CC1;
DT	01-OCT-2001 (TrEMBLrel. 18, Created)
DT	01-OCT-2001 (TrEMBRel. 18, Last sequence update)
DT	01-OCT-2001 (TrEMBRel. 18, Last annotation update)
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN	MLL5213.
OS	Rhizobium loti (Mesorhizobium loti).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC	Phyllobacteriaceae; Mesorhizobium.
OX	NCBI_TaxID=361;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=MAFF303099;
RC	MEDLINE=21082930; PubMed=11214968;
RX	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA	Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA	Tateuchi C., Yanada M., Tabata S.;
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT	Mesorhizobium loti.";
RL	DNA Res. 7:331-338(2000).
DR	EMBL; AP003006; BAB51700.1; -.
DR	InclPro; IPR001986; EPSP_syntase.
DR	Pfam; PF00275; EPSP_syntase; 1.
DR	ProDom; PD001867; EPSP_syntase; 1.
KW	Transferase; Complete proteome.
SO	SEQUENCE 452 AA; 47455 MW; 2B52983E3523B938 CRC64;

Query Match	77.9%	Score 1783;	DB 16;
Best Local Similarity	77.8%	Pred. No. 2.7e-100;	Length 452;
Matches 351;	Conservative 33;	Mismatches 63;	Indels 4;
Gaps 2;			

Db 121 GDASLSGRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPITYRVPMSAQVKS 180  
QY 180 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGGANITVETDADGVRTIRLEGRGKLTGQVI 239  
Db 181 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGGANITVETDADGVRTIRLEGRGKLTGQVI 240  
QY 240 DVPDGPSTAFPLVAALLVPGSDVTILNVLNMPRTGLTLTLOEMGADIEVINPRLAGGE 299  
Db 241 AVPGDPSAGFLVAALIVPGSDIYIENVLMNPTRTGLTLTLOEMGGQIDIINPRNAGGE 300  
QY 300 DVADLRVRSSTLKGVTVPEDRAPSMIDEXPIYLAFAAFAEGATVMNGLEELRVKESDRLS 359  
Db 301 DVADLRVRSSTLKGVTVPEDRAPSMIDEXPIYLAFAAFAEGATVMNGLEELRVKESDRLS 360  
QY 360 AVANGLKLVGDCDEGETSLVVRGPDGKGLG---NASGAAATVATHLDHRIAMSLFVNGLV 416  
Db 361 AVANGLKLVGDCDEGETSLVVRGPDGKGLG---NASGAAATVATHLDHRIAMSLFVNGLV 420  
QY 417 SENPVTVDATMIATSPFEMDLMAGLAKI 447  
Db 421 TEKPTIDQAMIAATSPFEMDLMAGLAKI 451

RESULT 4  
QY 421 Q9A2H2 PRELIMINARY; PRT; 443 AA.  
AC Q9A2H2;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE 3-PROSHOSHIMATE 1-CARBOXYVINYLTRANSFERASE.  
GN CC3589.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
Kolony J.F., Smith J., Craven M.B., Khouri H., Shetty J., Berry K.,  
Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
EMBL; AE006017; AAK25551.1; -;  
TIGR; CC3589; -;  
DR InterPro; IPR001986; EPPS\_synthase.  
DR Pfam; PF00275; EPPS\_synthase; 1.  
DR ProDom; PD001867; EPPS\_synthase; 1.  
DR PROSITE; PS00104; EPPS\_SYNTHASE\_1; UNKNOWN\_1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 443 AA; 46075 MW; 45E8A1463E10B6EC CRC64;

Query Match 49.6%; Score 1135; DB 16; Length 443;  
Best Local Similarity 55.8%; Pred. No. 4.5e-61;  
Matches 251; Conservative 45; Mismatches 138; Indels 16; Gaps 6;

QY 2 LHGASRRPATAKSGSLGTVRIPGDKSISHRSPFGLASGETRITGLLEGEDVINTGK 61  
Db 3 LAGLSAPGGA-----LRGIVRPGDKSISHRSPFGLASGETRITGLLEGEDVINTGK 57  
QY 62 ANOAMGARTKEG-DTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120  
Db 58 ANOAMGARTKEG-DTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 115  
QY 121 GDASLTKRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPITYRVPMSAQVKS 180

Db 116 GDQSLRGRPMGRVLEPLRQMGVQLKATPGDRMPITLHGPKHAAPITYRVPMSAQVKS 174  
QY 181 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGGANITVETDADGVRTIRLEGRGKLTGQ 237  
Db 175 VLLAGLNTPGITTVIEPIIMTRDHTKMLQGGANITVETDADGVRTIRLEGRGKLTGQ 234  
QY 238 VLDVPCDPSSTAFPLVAALLVPGSDVTILNVLNMPRTGLTLTLOEMGADIEVINPRLAG 297  
Db 235 HVAVPGDPSAGFLVAALIVPGSEVTVGVMLENEURTGLTLTLOEMGADLVISNVRAS 294  
QY 298 GEDVADLRVRSSTLKGVTVPEDRAPSMIDEXPIYLAFAAFAEGATVMNGLEELRVKESDR 357  
Db 295 GEEVGDTARYSOLKGVVVPEDRAPSMIDEXPIYLAFAAFAEGATVMNGLEELRVKESDR 354  
QY 358 LSANGLKLVGDCDEGETSLVVRGPDGKGLG---NASGAAATVATHLDHRIAMSLFVNGLV 417  
Db 355 ISLTANGLKLVGDCDEGETSLVVRGPDGKGLG---NASGAAATVATHLDHRIAMSLFVNGLV 410  
QY 418 ENPVTVDATMIATSPFEMDLMAGLAKI 447  
Db 411 QAEVAVDEPGMIATSPFEMDLMAGLAKI 440

RESULT 5  
QY 418 Q9H269 PRELIMINARY; PRT; 746 AA.  
AC Q9H269;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
DE STILL FRAMESHIFT 3-PROSHOSHIMATE 1-CARBOXYVINYLTRANSFERASE  
DE PREPHENATE DEHYDROGENASE.  
GN PA3164.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
"Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen.";  
Proc. Natl. Acad. Sci. U.S.A. 96:959-964(2000).  
EMBL; AE004740; RAG06552.1; -;  
DR InterPro; IPR001986; EPPS\_synthase.  
DR Pfam; PF00275; EPPS\_synthase; 1.  
DR ProDom; PD001867; EPPS\_synthase; 1.  
DR PROSITE; PS00104; EPPS\_SYNTHASE\_1; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 746 AA; 79320 MW; C2974B4BBF539E3A CRC64;

Query Match 43.2%; Score 989.5; DB 16; Length 746;  
Best Local Similarity 48.5%; Pred. No. 6e-52;  
Matches 214; Conservative 62; Mismatches 152; Indels 13; Gaps 3;

QY 12 ARKSSGSLGTVRIPGDKSISHRSPFGLASGETRITGLLEGEDVINTGKAMGARIR 71  
Db 317 APOGSSGSLGTVRIPGDKSISHRSPFGLASGETRITGLLEGEDVINTGKAMGARIR 376  
QY 72 -REGDTWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 130

Db	377	GPQNGRVTVHGVGLHGLKAPPGPIYVLSNGSGTSMRLLSGLLAAQPFDTLTGDSASLKRPM	436
Qy	131	GRVLNPLREMGVQKSGDGLPVTLRGPKTPTTIVRPMASQVKSASVALLAGLNTPGI	190
Db	437	NRVAKPLREMGAVIETGPEGRPTIRGQRLTGMHYDMPMASQVKSCLLGLAGLYAAGE	496
Qy	191	TTVIEPIWTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGGQVIDVPGDPSSTAF	250
Db	497	TSVTEPAPTRDHTERMLRGFGYPPVVEGS-----TAKVESGHKLSATHIEVPADISSAAF	551
Qy	251	PLVAALLVPGSDVTLNVLMPNTRTGLTLTQEMGADIEVINPRLAGGEDVADLRVRSST	310
Db	552	PLVASIAEGSELVQLHVGINPRTVGVTEILRLMGDDLSENQREVGEPVADIRVRSAR	611
Qy	311	LKGVTVPEDRAPSMIDIEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGV	370
Db	612	LKGIIDIPEDLVPLADEFPVLFVAAACAEGRVTLRGAELRVKESDRIQVMADGLKALGV	671
Qy	371	DCDEGETSLVVRGPDGKGLGNASCAAVATHLDRHIAVSFLVGLVSNPVTVDATMIA	430
Db	672	KAEPTPDGIVIEG-----GAGGGEVWAGHDHRIAMSFVSASLRASGPIRIHDCANVA	724
Qy	431	TSFPFMDLMAGLGAKIELSD 451	
Db	725	TSFPNLFALCAQTGIRVAVEN 745	
RESULT	6		
ID	Q9PB21	PRELIMINARY; PRT; 454 AA.	
AC	Q9PB21		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TRENBLrel. 19, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.		
GN	XF2324.		
OS	Xylella fastidiosa.		
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
OC	Xylella.		
OX	NCBI_TaxID=2371;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=9A5C;		
RX	MEDLINE=20365717; PubMed=10910347;		
RA	Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,		
RA	Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,		
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,		
RA	Bueno M.R.P., Camargo A.A., Canargo L.E.A., Carraro D.M., Carrier H.,		
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,		
RA	Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,		
RA	Faccinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,		
RA	Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,		
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,		
RA	Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,		
RA	Krieger J.E., Kuramae E., Laigret F., Lambais M.R., Leite L.C.C.,		
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,		
RA	Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.D.,		
RA	Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,		
RA	Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,		
RA	Moore D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,		
RA	Nanai A., Nobrega F.G., Nunes L.R., Oliveira M.A.,		
RA	de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,		
RA	Peixoto B.R., Perreira G.A.G., Perreira H.A. Jr., Pesquero J.B.,		
RA	Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,		
RA	de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,		
RA	da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,		
RA	da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,		
RA	de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshukano M.H.,		
RA	Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,		
RA	Zawo M.A., Zatz M., Meidanis J., Setubal J.C.;		
RT	"The genome sequence of the plant pathogen Xylella fastidiosa."		
RL	Nature 406:151-159(2000).		
CC	EMBL; AF04043; AF05123.1; -		

DR	InterPro: IPR001986; EPSP synthase.		
DR	Pfam: PF00375; EPSP synthase; 1.		
DR	ProDom: PD001867; EPSP synthase; 1.		
DR	PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.		
DR	PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.		
KW	Complete proteome.		
SEQ	SEQUENCE 454 AA; 48266 MW; 45CCF074E6C0BA57 CRC64;		
Query Match	39.9%; Score 912.5; DB 16; Length 454;		
Best Local Similarity	46.9%; Pred. No. 1.4e-47;		
Matches	206; Conservative 60; Mismatches 158; Indels 15; Gaps 4;		
Qy	12	ARKSSGLSGTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDEVINTGKAMQAMGARIR	71
Db	23	AHQGTPLHGVLSIPGDKSISHRVMAALADGTSRIDGFLEAEDCTCAETLARLGVRIE	82
Qy	72	KEGDTW-IIDGVNGGLLAPLDFGNATGCRLTGLVGVYDFDSTFTGDSASLTKRPM	130
Db	83	TPLSTQIRIVHGVGVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVDASLSKRP	142
Qy	131	GRVLNPLREMGVQKSGDGLPVTLRGPKTPTTIVRPMASQVKSASVALLAGLNTPGI	190
Db	143	RRVTDPLSQMGARITSDGTPPLRIYGGQLLHGIDFISPVASQIKSAVLLAGLYARNE	202
Qy	191	TTVIEPIWTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGGQVIDVPGDPSSTAF	250
Db	203	TVREPHPTRDYTERMLTAFGVLDIVSTGC-----ARLGGQRLCATDITIPADFSAAF	257
Qy	251	PLVAALLVPGSDVTLNVLMPNTRTGLTLTQEMGADIEVINPRLAGGEDVADLRVRSST	310
Db	258	YLVAASVTPGSDITLRAVGLNPRRIGLITVLRMGANIVESNRHEQGEPPVVDLRVYAP	317
Qy	311	LKGVTVPEDRAPSMIDIEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGV	370
Db	318	LOGTRVPEDLVADMTIDEFPALFVAAAAGQTVVSGAAELRVKESDRLLAAVMTGLRVLG	377
Qy	371	DCDEGETSLVVRGPDGKGLGNASCAAVATHLDRHIAVSFLVGLVSNPVTVDATMIA	430
Db	378	QVDETFADGATGGPIGHGTINSHG-----DHRIMAFSIAGLSYSTVRIEDVANVA	430
Qy	431	TSFPFMDL--MAGLGAKI 447	
Db	431	TSFPDYETLARSAGFGLEV 449	
RESULT	7		
ID	Q9KCA6	PRELIMINARY; PRT; 431 AA.	
AC	Q9KCA6		
DT	01-OCT-2000 (TRENBLrel. 15, Created)		
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)		
DT	01-OCT-2001 (TRENBLrel. 18, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS).		
GN	AROF OR BH1667		
OS	Bacillus halodurans.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/staphylococcus group; Bacillus.		
OX	NCBI_TaxID=86665;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=C-125 / JCM 9153;		
RX	MEDLINE=20512582; PubMed=11058132;		
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,		
RA	Fuji F., Hirano C., Nakamura Y., Ogawara N., Kuhara S.,		
RA	Horikoshi K.;		
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus		
RT	halodurans and genomic sequence comparison with Bacillus subtilis."		
RL	Nucleic Acids Res. 28:4317-4331(2000).		
CC	-1- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -		
	ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.		



CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE  
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -!- SIMILARITY: BELONGS TO THE EFSP SYNTHASE FAMILY.  
DR EMBL; AP001512; BAB05386.1; -.  
DR InterPro; IPR001986; EFSP\_synthase.  
DR Pfam; PF00275; EFSP\_synthase; 1.  
DR ProDom; PD001867; EFSP\_synthase; 1.  
DR PROSITE; PS00104; EFSP\_SYNTHASE\_1; 1.  
DR PROSITE; PS00885; EFSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; Complete proteome; Transferase.  
SQ SEQUENCE 431 AA; 45485 MW; 12F4FBE7BA0743D CRC64;

Query Match 38.6%; Score 884; DB 16; Length 431;  
Best Local Similarity 44.2%; Pred. No. 6.9e-46;  
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;  
QY 15 SSGLSGTVIRPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIRKEG 74  
10 AKGLGKTIKVPDGSISHRAVMFEGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69  
75 DTWIIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134  
70 ERVTVKGGWDLREPSDILDVNGSGTTTTLTGLTSLTFPHSVIIGDESIGKRPKRVY 129  
QY 135 NPLRMGVQVKSED-GDRLPVTLRGPKTPTPTIRYVPMAAQAQVKSALLAGLNTPGITTV 193  
DB 130 EPLKSMGAQIDGRDGNLTPLSIRGQL-KGIDFISPVASAKMSAILLAGURAGKTSV 188  
QY 194 IEPIMTRDHTKMLQGFANLTVETDADGVTRIRLEGKGLTQGVYDVPDSSFAFLV 253  
DB 189 TEPAKTRDHTERMLEAFGNI----EKDGL-TVSIEGGQMLTGQVVPDGISSAAFLV 243  
QY 254 AALLVPGSDVTLNVLMPTRTGLTILTOEMGADIEVINPRLAGGEDVADLVRSTLKG 313  
DB 244 AGAMVPHSRITLTNVGINPTFRAGILEVLKQMGATLAMENRVQGGEPVADLTETISVLOG 303  
QY 314 VTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNCVDCD 373  
DB 304 VEIGDIIIPRLIDEIPIITAVLATQASGRVTKDAELKVKETNRIDTVVSELTKLKGASTH 363  
QY 374 EGETSLVVRGPDGKGLGNASGAAVATHLDHRIAMSLVNGLVNPNVTVDDATMATSF 433  
DB 364 ATDDGMIEGTPPLAG-----GVTVSSHCHDRIGMAMATAALLAEKPVTEGTEALNSV 418  
QY 434 PEFMD 438  
419 PSFFD 423

RESULT 8  
Q99283 ID Q99283 PRELIMINARY; PRT; 430 AA.  
AC Q99283;  
DT 01-JUN-2001 (TremBLrel. 17, Created)  
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXIVINYLTRANSFERASE  
DE (EC 2.5.1.19).  
GN AROA OR SPY1352.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;  
RX MEDLINE=21192684; PubMed=11296296;  
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
RA Primeaux C., Seaton S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).  
DR EMBL; AE006573; AAK34180.1; -.  
DR InterPro; IPR001986; EFSP\_synthase.  
DR Pfam; PF00275; EFSP\_synthase; 1.  
DR ProDom; PD001867; EFSP\_synthase; 1.  
DR PROSITE; PS00104; EFSP\_SYNTHASE\_1; 1.  
DR PROSITE; PS00885; EFSP\_SYNTHASE\_2; UNKNOWN\_1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 430 AA; 46692 MW; 3273C6B39020FB61 CRC64;  
Query Match 35.8%; Score 820; DB 16; Length 430;  
Best Local Similarity 43.8%; Pred. No. 5.1e-42;  
Matches 189; Conservative 76; Mismatches 141; Indels 26; Gaps 9;  
QY 13 RKSSG-LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTKAMQAMGARIR 71  
DB 7 RTNAGPLQCTIQVPGDKSISHRAVLGAVAKGTRVKGGLKGLDVLSTIOAFNGLGVRIE 66  
QY 72 KEGDTWIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFIGDASLTKRPMG 131  
DB 67 EKDDQLVIEGQFGQLNAPCQTLNMGNSGTSMRLIAGLLAGQPFVSKMIGDESLSKRPMD 126  
QY 132 RVLNPLRMGVQVKSP-DGRLPVTLRGPKTPTPTIRYVPMAAQAQVKSALLAGLNTPGI 190  
DB 127 RIVYPLKQMGVEISGETDRQFPPLQOGNRLQPIYTLPISSAQVKSALLAALQAKGT 186  
QY 191 TVIPEIMTRDHTKMLQGFANLTVETDADGVTRIRLEGKGLTQGVYDVPDPSSTAF 250  
DB 187 TQVEKEITRNTEEMIQQFGRLIV---DGKR-ITLVGPQQLTAQETVPGDISSAAF 241  
QY 251 PLVAALLVPGSDVTLNVLMPTRTGLTILTOEMGADI--EYINPRLAGGEDVADLVRVS 308  
DB 242 WLVAGLIIPGSELLKKNVGNPTRTGILEVWKMGAIQIYVEDMNRK-----EQVTSIRVY 297  
QY 309 STLKGVTPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGKLKN 368  
DB 298 SNMKGTTISGLIIPRLIDEIPIITAVLATQASGTCTCKDAQELVETDRIOVVTDILNSM 357  
QY 369 GYDCDEGETSLVVRGPDG---KGLGNASGAAVATHLDHRIAMSLVNG-LVSENPVTVD 424  
DB 358 GAN-----IKATADGMIKGPVLYGANTSTYGDHRIGMTAIAALLVKQGQVHLD 408  
QY 425 DATMTATSPPEP 436  
DB 409 KEEAIMTSYPTF 420  
RESULT 9  
Q92A85 ID Q92A85 PRELIMINARY; PRT; 428 AA.  
AC Q92A85;  
DT 01-DEC-2001 (TremBLrel. 19, Created)  
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE AROE PROTEIN  
GN AROE OR LIN2037.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CLIP 11262 / SEROVAR 6A;  
RX PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kuhn M., Kunst F., Kurapkat G.,  
RA Madieno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,  
RA "Comparative genomics of *Listeria* species.",  
RL Science 294:849-852(2001).  
RL ENBL; AL596170; CAC97267.1; -.  
DR ListList; LIN02037; -.  
KW Complete proteome.  
SQ SEQUENCE 428 AA; 45994 MW; 157B48C091A68FEB CRC64;

Query Match 35.68; Score 815.5; DB 16; Length 428;  
Best Local Similarity 40.8%; Pred. No. 9.5e-42;  
Matches 173; Conservative 83; Mismatches 155; Indels 13; Gaps

QY 17 GLSGTVRVPDKSISHRSMFGVGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT 76  
DB 9 GLVEITVPGDKSMHSRSMIFGAIEGKTVIRHFLRADDCGLTAKAFKALGVKIEETDEE 68  
QY 77 WIIDGVNGGLLAPADPFDGNAATGCRLTWGLVGVYDFDSTFIGDASLTKEPGRVLP 136  
DB 69 IIVHGTSGDLKQABGPLDIGNSGTTIRLMGILAGRDFDTVILGDSTAKRPMRVMLP 128  
QY 137 LREMGSVQVKSDEGDRLL-PVTLRGPKTPITTVRVPMASQAQVSAVLLAGLNTPGTITVIE 195  
DB 129 LOEMCAKWHKRDGSEFAPIISIGNSLKRMHEYHMPVSAQVKSATFAALQAEGETIIE 188  
QY 196 PIMTRDHTKMLQGGANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSSTAPLVA 255  
DB 189 KEKTRDHTHEMTIRQFGG---EIMDGL-TIRVKGQKFIGQEMTPGVGVSSAAFFVAG 243  
QY 256 LLVPGSDVTLNVLNPNRTGILTLIQEMGADIEVINPRLAGGEDVADLRVRSSTLKGYT 315  
DB 244 LITPGESELTWGLNPNRTGTFIDVVEQMGSSILVWKDSSRSTGKLAGTVVKSSELKGT 303  
QY 316 VPEDRAPSMIDEPILAAVAFAEGATVNGLEELRYKESDRLSAVANGKLNGVDCDEG 375  
DB 304 IGGDIIPRLDIEPVIALLATQAEGTITIKDAELKVKETNRIDAVANELNKGADITPT 363  
QY 376 ETLVYVRGPDGKGLNAGAAVATHLDHRIAMSEFLVMG-LVSENPVTDDATMIATSP 434  
DB 364 EGLIIRGTP-----LHAANVTSYGDHRIGMMLQIALLVEDGDVELDRAEAVSVYP 417  
QY 435 EFMD 438  
DB 418 TFFE 421

RESULT 10  
ANY6  
QY 09ANY6 PRELIMINARY; PRT; 428 AA.  
AC Q9ANY6;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXIVINYLTRANSFERASE (EC 2.5.1.19).  
GN AROA.  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
OC Enterococcus.  
OX NCBI\_TaxId=1351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OG1RF;  
RA Huycke M.M., Shepard L., Joyce W., Wise P., Moore D.R., Gilmore M.S.;  
RT "Forme fruste respiration by *Enterococcus faecalis* produces extracellular superoxide".  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AP318277; AAG53678.1; -.  
DR InPro; IPR001986; EPSP\_synase.  
DR Pfam; PF00275; EPSP\_synase; 1.  
DR ProDom; PD001867; EPSP\_synase; 1.  
DR PROSITE; PS00104; EPSP\_SYNTHASE\_1; UNKNOWN\_1.

DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; UNKNOWN\_1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 432 AA; 47068 MW; A42102057AD15C72 CRC64;

Query Match 26.7%; Score 610; DB 16; Length 432;  
Best Local Similarity 34.0%; Pred. No. 2.7e-29;  
Matches 146; Conservative 80; Mismatches 181; Indels 22; Gaps 6;

QY 15 SSGLSGTVPIDPKSHRSFMEFGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74  
DB 10 SGPLKEIEVPDGKSTHRAIMLASLAEGTSNIYKPLLEGDCRRMDIFRLGVDIKDE 69

QY 75 DTWIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFGDASLTKRPMGRVL 134  
DB 70 DKLNVNSPGYKAFKTPHQVLYTNGSGTTRLLAGLLSGLGIESVLSDYSIGKRPMDRVL 129

QY 135 NPLREMGVOVKSEGDRLPVTIRGPKTPITYRVPMAOQKSAVLLAGLNTPTGTTVI 194  
DB 130 RPLKMDANIEGIEDNYTLPIK-PSVIKGINYQMEVASAOQKSAIFASLFSNDTTVIK 188

QY 195 EPIMTRDHTKMLQGF-----GANLTVDADGVRTIRLEGRGKLTGQVTDVPGDPSSA 249  
DB 189 ELDVSRNHTETFRFNIEPIERISITTPDAIQIKPAD-----FHVPGDISAA 240

QY 250 FPLVAALLVPGSDVTILNLMNPTRLTLOEMGADIEVINPRLAGGEDVADLRV-S 308  
DB 241 FFIVAALITPESDVTIHNVGINPTSGIIDIIVKMGNGIOLFEN-QTTGAEPASIRIQYT 299

QY 309 STLKGVTPEDRAPSMIDEPYILAVANAFAEGATVNGLEELRVKESDRLSAVANGLKN 368  
DB 300 PMLQPTITIEGVLPAKIDELPVALICTOAVGTSTIKDAEELKVKETNRIDTADMLNLL 359

QY 369 GYDCDEGETSLVVRGPRDGKGLGNAGAAVATHLDRIAMSLVGLVSENPTVVDATM 428  
DB 360 GFELQPTNDGLIHPSE-----FKTNAVVDLSLTHRIGMMLAVASLLSEPKIKQFDA 413

QY 429 IATSPPEFM 437  
DB 414 VNVSEFGFL 422

RESULT 12  
QYRH28 PRELIMINARY; PRT: 207 AA.  
QYRH28; MEDLINE=20504483; PubMed=11016950;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE 5-ENOLPYRUVLSHKIMATE 3-P SYNTHASE (FRAGMENT).  
GN AROF.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JM300;  
RX MEDLINE=92013931; PubMed=1919506;  
RA Fischer R.S., Zhao G., Jensen R.A.;  
RT "Cloning, sequencing, and expression of the P-protein gene (phea) of  
RT Pseudomonas stutzeri in Escherichia coli: implications for  
RT evolutionary relationships in phenylalanine biosynthesis";  
RL J. Gen. Microbiol. 137:1293-1301(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JM300;  
RX MEDLINE=99298294; PubMed=10368439;  
RA Xie G., Bonner C.A., Jensen R.A.;  
RT "A probable mixed-function supraoperon in Pseudomonas exhibits gene

RT organization features of both intergenomic conservation and gene  
RT shuffling";  
RL J. Mol. Evol. 49:108-121(1999).  
DR EMBL: AF038578; AAD47363.1; -.  
DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; UNKNOWN\_1.  
FT NON\_TER 207  
SQ SEQUENCE 207 AA; 21780 MW; 7231191C72A21D6B CRC64;

Query Match 20.6%; Score 470.5; DB 2; Length 207;  
Best Local Similarity 51.5%; Pred. No. 2.7e-21;  
Matches 103; Conservative 22; Mismatches 70; Indels 5; Gaps 2;

QY 12 ARKSSGLSGTVPIDPKSHRSFMEFGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71  
DB 10 ANPGSISLGTOLRYPGDKSISHRSIMLSLAEGTTEXEGFLEGEDALATXQAFRXMGVVI- 68

QY 72 KEG---DTWIIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFGDASLTKR 128  
DB 69 -EGPHQGRVTVHGVGLHQAQPPGPIYLGNSGTSMRLLAGLLAAQPFDTTFLSGDASLTKR 127

QY 129 PMGRVNLPLREMGVOVKSEGDRLPVTIRGPKTPITYRVPMAOQKSAVLLAGLNTPT 188  
DB 128 PMNRVAKPLREMGAVIETAAEGRPLTIRGGKKLSGHHYDMPMASAOQKSCLLAGLYAA 187

QY 189 GITTVEIPINTRDHTEKMLQ 208  
DB 188 GKTSVTEPAPTRDHTERMQ 207

RESULT 13  
QYHQC1 PRELIMINARY; PRT: 439 AA.  
ID Q9HQC1  
AC Q9HQC1; (TREMBLrel. 16, Created)  
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.  
GN PSC OR VNG1232G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weiti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.N., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005049; AAG19594.1; -.  
DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; UNKNOWN\_1.  
DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; UNKNOWN\_1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 439 AA; 44333 MW; 48CAD75A1F0ECB89 CRC64;

Query Match 20.6%; Score 470.5; DB 17; Length 439;  
Best Local Similarity 31.3%; Pred. No. 7.6e-21;  
Matches 139; Conservative 76; Mismatches 190; Indels 39; Gaps 13;

Qy	16	SGLSGTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDEVINTCKAMQAMCARIRKEGD	75
Db	17	SKVRGRARAPPSKSYTHRALAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG	76
Qy	76	TWIDGVNGGGLLAPEAPLDFGNATGRLTMGLGVGVDFDSTFIGDASLTGKRMGRVLN	135
Db	77	DWVTVFGSFPAL-PAVIDCANSCTMRLVTAATAADAGTTVLTGDESRLRPHGPLLD	135
Qy	136	PLREMGVQVKSDGD-RLPVTLRGPKPTPTITYRVP-MASAQVKSAVLLAGLNT-----P	188
Db	136	ALSGLGTARSTRGNGQAPLVDPVSGGSA--LPGDVSSQFVTALLMAGAVTETGIET	193
Qy	189	GITTVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRG-----KLTGOVIDVPGD	244
Db	194	DLTTELSAPYVDITLDVDAFGVAGS-ET-ANGYRV-----RGQAYAPSGAEYAVPGD	246
Qy	245	PSSTAFPLVAALLVPGSDVTTILVLMNPTRTG---LILTLOEMGADIETVNPRLAGGEDV	301
Db	247	FSSASYLLAAGALAAADGAAVVEGMHPSAQGDAAIVDVLRMGADID-----WDTES	299
Qy	302	ADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEAGATVMNGLEELRVKESDLSAV	361
Db	300	GVTIVQSRSELSGVEGVADTDPDL---PTIAVLGAADGTTTRITDAEHVRYKETDRVAA	356
Qy	362	ANGLKLVGDCDEGETSLVVRGPDGKGLGNASGAATHLDHRIAMSFVLMGLVSENPV	421
Db	357	AESLSKLGSVEERDELVVRG-----GDTLSGASVDGGRDHLVLMALAVAGLVADGET	411
Qy	422	TVDDATMIATSPFPEMDLMAGLGA	445
Db	412	TIAGSEHVDVSPDFFEVLGLGA	435
RESULT 14			
Q9L213	PRELIMINARY; PRT; 440 AA.		
ID	Q9L213		
AC	Q9L213		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.		
GN	ARO.		
OS	Streptomyces coelicolor.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.		
OX	NCBI_TaxID=1902;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RL	STRAIN=A3(2);		
RL	Brown S.P., Harris D.;		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3(2);		
RA	Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3(2);		
RX	MEDLINE=97000351; PubMed=8843436;		
RA	Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,		
RA	Kinashi H., Hopwood D.A.;		
RT	"A set of ordered cosmids and a detailed genetic and physical map for		
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";		
RL	Mol. Microbiol. 21:77-96(1996).		
DR	EMBL; ALI38598; CAB71266.1; -		
DR	InterPro; IPR001986; EPSP_synthase.		
DR	InterPro; IPR000408; RCC1.		
DR	Pfam; PF00275; EPSP_synthase; 1.		
DR	ProDom; PD001867; EPSP_synthase; 1.		
DR	PROSITE; PS00885; EPSP_SYNTHASE.2; 1.		
DR	PROSITE; PS00626; RCC1_2; UNKNOWN_1.		
KW	Transferase.		

Qy	11	TARKSSGTSCTVRIPGDKSIHRSFMFGGLASGETRITGLLEGEDEVINTGKAMQAMG---	67
Db	15	TITLTGDETVRVLGSKSYNRYLAISLGSQETVIDNALLSDDTVFYSRAIETFGHVT	74
Qy	68	-----ARIRKEGDTWIIDGVNGGGLLAPEAPLDFGNATGRLTMGLGVGVDFDSTFI	120
Db	75	CDIDHATARIR-----VTPTGR-PMRAPSEDFVGGAGCTPLRFLISMAGHADGTTIIT	126
Qy	121	GDASLTGRPGRVLNPLREMGVQVKSDGD-RLPVTLRGPKPTPTITYRVPMSAQVKA	179
Db	127	GNARQMRQMGDLKALPALGVDAVATVRNGSPVVRVGGSGKGGATSGISGAVSSQTS	186
Qy	180	VLLAGLNTPGIT--TVIEPIMRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTGQ	237
Db	187	LIINALRAQTDTEITISDDLVSQPYVEMTLAGL-AEMGVSVDRDGYRFTVPSGQARG	245
Qy	238	VIDVPGDPSSTAFPLVA-----LLVPGSDVTTILVLMNPTRTGLILTLOEMGADIEVI	291
Db	246	QVTVEPDASGMSYFLAAAILQSRVVIIPG-----IGAGSHOGDVHLVQALERMGRTEV	299
Qy	292	NPLRAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEAGATVMNGLEELR	351
Db	300	-----GDD--SIIVTGGPLRGIDIDMEAMPDVV---PSLAIVAAAYABGTTITNIASLR	348
Qy	352	VKESDRLSAVANGKLKLVGDCDEGETSLVVRGPDGKGLGNASGAATHLDHRIAMSF	411
Db	349	VKECDRIAATTELKMGIDVEEHSADAMYITG-----GTPHGAVIDTYDDHRIAMTFA	401
Qy	412	VNGLVSENPVTVDDATMIATSPFPEMDLMAGLGA	448
Db	402	IGGLRTEG-VWIKDPCVAKSPFAPFWQITDLTHPDL	437
RESULT 15			
Q9KR80	PRELIMINARY; PRT; 426 AA.		
ID	Q9KR80		
AC	Q9KR80		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-		
DE	ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)		
DE	(EPSPS).		
GN	VC1732.		
OS	Vibrio cholerae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
OX	NCBI_TaxID=666;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;		
RX	MEDLINE=20406833; PubMed=10952301;		
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,		
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,		
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,		
RA	Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,		
RA	McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,		
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,		
RA	Fraser C.M.;		
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio		
RT	cholerae".		
RL	Nature 406:477-483(2000).		
CC	-!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -		
CC	ORTHOPOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.		
CC	-!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE		
CC	AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).		
CC	-!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.		

Qy	18	28; Score 416.5; DB 2; Length 440;	
Qy	28	0%; Pred. No. 1.4e-17;	
Qy	128	Conservative 78; Mismatches 198; Indels 53; Gaps 13;	

Search completed: August 15, 2002, 14:05:44  
Job time: 413 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:47 ; Search time 56.25 Seconds  
(without alignments)  
898.464 Million cell updates/sec

Title: US-09-464-099A-70  
Perfect score: 2288  
Sequence: 1 MLHGASRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Genesec-052802.\*

```

1: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	13 AAR26449	CP4-EPSPS. Synthe
2	2288	100.0	455	19 AAW34685	CP4-EPSPS protein.
3	2288	100.0	527	22 AAM52214	EPSPS SEQ ID NO 3.
4	2282	99.7	455	13 AAR22300	Class II EPSPS enz
5	2282	99.7	455	18 AAW34683	Class II EPSP synt
6	2282	99.7	455	18 AAW24474	Class II EPSPS for
7	2282	99.7	455	19 AAW71609	Agrobacterium sp.
8	2282	99.7	455	22 AAE05053	Agrobacterium sp.
9	1900.5	83.1	449	13 AAR22301	Class II EPSPS enz
10	1900.5	83.1	449	13 AAR22302	Class II EPSPS enz
11	1900.5	83.1	449	18 AAW34684	Class II EPSP synt

12	1900.5	83.1	449	18 AAW34685	Class II EPSP synt
13	1900.5	83.1	449	18 AAW24479	Class II EPSPS use
14	1900.5	83.1	449	18 AAW24480	Class II EPSPS use
15	1900.5	83.1	449	19 AAW71611	Pseudomonas sp. st
16	1900.5	83.1	449	19 AAW71610	Achromobacter sp.
17	1900.5	83.1	449	22 AAE05054	Achromobacter sp.
18	1900.5	83.1	449	22 AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18 AAW34688	Class II EPSP synt
20	867.5	37.9	447	18 AAW24487	Class II EPSPS for
21	867.5	37.9	447	19 AAW71619	Synechocystis sp.
22	867.5	37.9	447	22 AAE05072	Synechocystis sp.
23	806	35.2	443	18 AAW34689	Class II EPSP synt
24	806	35.2	443	18 AAW24488	Class II EPSPS for
25	806	35.2	443	19 AAW71620	Dichelobacter nodo
26	806	35.2	443	22 AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22 AAB48177	S. pneumoniae aroA
28	803.5	35.1	427	22 AAB37094	Streptococcus pneu
29	803.5	35.1	427	22 AAB48500	Streptococcus pneu
30	798.5	34.9	427	20 AAW97388	5-enolpyruvylshiki
31	798.5	34.9	427	21 AAY56504	Streptococcus pneu
32	798.5	34.9	427	22 AAB31158	A 5-enolpyruvylsh
33	785	34.3	428	18 AAW34686	Class II EPSP synt
34	785	34.3	428	18 AAW24481	Class II EPSPS use
35	785	34.3	428	19 AAW71617	Bacillus subtilis
36	785	34.3	428	22 AAE05070	Bacillus subtilis
37	704.5	30.8	415	20 AAW97389	protein encoded by
38	704.5	30.8	415	21 AAY56505	Streptococcus pneu
39	704.5	30.8	415	22 AAB31159	A 5-enolpyruvylsh
40	704.5	30.8	415	22 AAB48178	S. pneumoniae aroA
41	704.5	30.8	415	22 AAB37095	Streptococcus pneu
42	704.5	30.8	415	22 AAB48501	Streptococcus pneu
43	577	25.2	430	18 AAW34687	protein encoded by
44	577	25.2	430	18 AAW24482	Class II EPSP synt
45	577	25.2	430	19 AAW71618	Class II EPSPS use
					Staphylococcus aur

#### ALIGNMENTS

```

RESULT 1
AAR26449
ID AAR26449 standard; Protein; 455 AA.
XX
AC AAR26449;
XX
DT 28-JAN-1993 (first entry)
XX
DE CP4-EPSPS.
XX
KW PMON1030; CTP2; CP4; EPSPS; chloroplast transit peptide;
KW 5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;
KW Pseudomonas chloroaphis 6G5.
XX
OS Synthetic.
XX
PN WO9212249-A.
XX
PD 23-JUL-1992.
XX
PF 17-DEC-1991; 91WO-US09437.
XX
PR 26-DEC-1990; 90US-0632440.
XX
PA (MONS ) MONSANTO CO.
XX
PI Kishore GM, Klee HJ;
XX
DR WPI; 1992-284334/34.
XX
N-PSDB; AAQ27201.
XX
Delaying fruit ripening and senescence in plants - by controlling
ethylene prodn., pref. by expression of
1-amino:cyclopropane-1-carboxylic acid deaminase

```





AC AAM52214;  
 XX 12-FEB-2002 (first entry)  
 XX EPSPS SEQ ID NO 3.  
 DE  
 XX Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;  
 KW 5-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;  
 KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;  
 KW Chloroplast transit peptide.  
 XX  
 OS Petunia hybrida.  
 OS Agrobacterium sp.  
 XX  
 PN JP2001190168-A.  
 XX  
 PD 17-JUL-2001.  
 XX  
 XX 27-OCT-2000; 2000JP-0328811.  
 XX  
 XX 29-OCT-1999; 99JP-0310244.  
 XX  
 XX (SUMO ) SUMITOMO CHEM CO LTD.  
 XX  
 XX WPI; 2001-605307/69.  
 DR N-PSDB; ABA02854.  
 XX  
 XX New herbicide-resistant plant -  
 PT  
 XX  
 PS Examples; Page 46-48; 66pp; Japanese.  
 XX  
 CC The invention relates to a transgenic plant which shows resistance to a  
 CC herbicide in an amount inhibiting natural  
 CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant  
 CC and having at least one enzymatic activity selected from:  
 CC (1) EPSPS activity different from natural EPSPS activity of the plant or  
 CC (2) glyphosate oxidoreductase (GOX) activity different from the natural  
 CC GOX activity of the plant; and in which a gene encoding a protein having  
 CC the following properties: (a) combines specifically with a substance  
 CC participating to the herbicidal activity a herbicide of  
 CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no  
 CC denaturing activity on a substance to which said protein combines  
 CC specifically; and (c) contains substantially no framework region of the  
 CC variable region of immunoglobulin. The present sequence is that of the  
 CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium  
 CC sp. strain CP4 EPSPS gene.  
 XX  
 XX Sequence 527 AA;  
 Query Match 100.0%; Score 2288; DB 22; Length 527;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-181;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLHGASSRPATARKSSGLSGTVPRIKSGTSHRSFPGGLASGETRITGLLEGEDVINTG 60  
 DB 73 mlhgassrpataarkssglsgtvpripksgtshrsfmgfpglasgetritgllegedvintg 132  
 QY 61 KAMQANGARIRKEGDTWIIDGVNGGGLLAPEAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120  
 DB 133 kamqangarirkegdtwiiidgvngngllapeapldfgnaatgcrlltmglvgvydfdstfi 192  
 QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPPIYRVPMASQAQVKS AV 180  
 DB 193 gdasltkrmpgrvnlplremgvqvksedgdlrplvtlrgpktptpityrvpmasaqvksav 252  
 QY 181 LLAGLNTPGITTTVEIPIMTRDHTKMLQCGFANLTVETDADGVRTIRLEGRGLTGOVID 240  
 DB 253 llaglnptgittveipimtrdhtekmlqgfanltvetdaddgvrtirlegrgkltgqv id 312  
 QY 241 VPGDPSSTAFPLVAALLVPGSDVTILNLVLMNPRTTGLILFLQEMGADIIEVINPRLAGGED 300  
 DB 313 vpgdpsstafplvaallvpgsdvtiinlvlnmnprrttrtglilflqemgadiev inprlagged 372

QY 301 VADLRVRSSTLKGVTPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
 DB 373 vadlrvrstlkgtvpedrapsmideyplilavaaafegatvmngleelrvkesdr lsa 432  
 QY 361 VANGKLINGVDCDEGTSLVVRGRPDGKGLGNASGAATAVTHLDRHRIAMSLVMGLVSEN P 420  
 DB 433 vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamslvmnglvsenp 492  
 QY 421 VTVDATMIATSPPEFMDLMAGLGAKIELUSDTRKAA 455  
 DB 493 vtvdattmiatspefmdlmaglgakielelsdtkaa 527  
 RESULT 4  
 AAR22300  
 ID AAR22300 standard; Protein; 455 AA.  
 XX  
 AC AAR22300;  
 XX  
 DT 03-AUG-1992 (first entry)  
 XX  
 DE Class II EPSPS enzyme.  
 XX  
 KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;  
 KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.  
 XX  
 OS Agrobacterium sp. strain CP4.  
 XX  
 PN WO9204449-A.  
 XX  
 PD 19-MAR-1992.  
 XX  
 PF 28-AUG-1991; 91WO-US06148.  
 XX  
 PR 31-AUG-1990; 90US-0576537.  
 XX  
 PA (MONS ) MONSANTO CO.  
 XX  
 PI Barry GF, Kishore GM, Padgett SR;  
 XX WPI; 1992-114356/14.  
 DR  
 XX DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate  
 PT synthase - for producing plants and bacteria tolerant to  
 PT glyphosate herbicides  
 XX  
 PS Disclosure; Fig 3; 148pp; English.  
 XX  
 CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3  
 CC phosphate synthase enzyme (EPSPS) of Agrobacterium sp. strain CP4  
 CC It is used to create glyphosate resistant plants or seeds which  
 CC can be planted in a field of crops to selectively control weeds.  
 CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,  
 CC tobacco and alfalfa. This provides a cost effective, environmentally  
 CC compatible weed control device. See also AAR22301 and AAR22302.  
 XX  
 XX Sequence 455 AA;  
 Query Match 99.7%; Score 2282; DB 13; Length 455;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-180;  
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLHGASSRPATARKSSGLSGTVPRIKSGTSHRSFPGGLASGETRITGLLEGEDVINTG 60  
 DB 1 mshgassrpataarkssglsgtvpripksgtshrsfmgfpglasgetritgllegedvintg 60  
 QY 61 KAMQANGARIRKEGDTWIIDGVNGGGLLAPEAPLDFGNAATGCRLLTMGLVGVYDFDSTFI 120  
 DB 61 kamqangarirkegdtwiiidgvngngllapeapldfgnaatgcrlltmglvgvydfdstfi 120  
 QY 121 GDASLTKRPMGRVNLPLREMGVQVKSEDDGRLPVTLRGPKTPPIYRVPMASQAQVKS AV 180

Db	121	gdsaltrkpmgrvlnplremngvqksedgdrplvtlrgpktpitpityrvpmasaqvksav	180
Qy	181	LLAGLNTPGITVTIEPTWTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGQVID	240
Db	181	llaglnpvgitvtiepimtrdhtekmlqgfganltvetdadvrtirlegrkltgqvld	240
Qy	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPRTTGILILQEMGADIEVINPRLAGGED	300
Db	241	vpgdpsstafplvaallvpgsdvtilnvlmptrtgililtqemgadievlnprlagged	300
Qy	301	VADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA	360
Db	301	vadlrvrstlkgvtvpedrapsmidpilaavaafaegatvmngleelrvkesdrlsa	360
Qy	361	VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSFVLVGLVSEN	420
Db	361	vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamsflvmglvsenp	420
Qy	421	VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA	455
Db	421	vtvdatmiatsfpefmdlmaglakieisdtkaa	455
RESULT 5			
AAW34683			
ID	AAW34683 standard; Protein; 455 AA.		
AC	AAW34683;		
XX			
DT	17-FEB-1998 (first entry)		
XX	Class II EPSP synthase (EPSPS) from Agrobacterium sp. strain CP4.		
KW	5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;		
KW	shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;		
KW	5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;		
KW	glyphosate resistance gene; glyphosate-tolerance; promoter.		
XX	Agrobacterium sp. strain CP4.		
XX	Key	Location/Qualifiers	
FT	Region	200..204	
FT		/label= characteristic_region	
FT		/note= "see AAW34690"	
FT	Region	26..29	
FT		/label= characteristic_region	
FT		/note= "see AAW34691"	
FT	Region	173..177	
FT		/label= characteristic_region	
FT		/note= "see AAW34692"	
FT	Region	271..274	
FT		/label= characteristic_region	
FT		/note= "see AAW34693"	
XX	US5627061-A.		
PN			
XX			
PD	06-MAY-1997.		
XX			
PF	07-JUN-1995; 95US-0476008.		
XX			
PR	13-SEP-1994; 94US-0306063.		
PR	31-AUG-1990; 90US-0576537.		
PR	28-AUG-1991; 91US-0749611.		
XX			
PA	(MONS ) MONSANTO CO.		
XX			
PI	Barry GF, Kishore GM, Padgett SR, Stallings WC;		
XX			
DR	WPI; 1997-271315/24.		
DR	N-PSDB; AAT93788.		
XX	Production of glyphosate-herbicide tolerant plants - using DNA		
PT			

encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase enzyme(s)			
Claim 7; Fig 3; 151pp; English.			
AAW34683-89 represent a new class of glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel EPSPS enzymes have little homology with known Class I EPSPS enzymes, and belong to a new class, Class II. The present sequence is from Agrobacterium sp. strain CP4. The EPSPS enzymes are part of the shikimic acid pathway, which leads to the biosynthesis of aromatic compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and is inhibited by the 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the herbicide glyphosate. It would be useful to produce transgenic crops containing glyphosate resistance genes so that glyphosate-containing herbicides can be applied to selectively kill weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that when introduced into a plant, the plant is made glyphosate-tolerant, and EPSPS enzyme activity is not affected. These class II EPSPS enzymes are fused to a chloroplast transit peptide to target the protein into the chloroplast, which is the site for the shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant under the control of a promoter such as figwort mosaic virus promoter or the cauliflower mosaic virus promoter, so that expression is enhanced.			
XX	Sequence 455 AA:		
SQ			
	Query Match	99.7%; Score 2282; DB 18; Length 455;	
	Best Local Similarity	99.8%; Pred. No. 1.1e-180;	
	Matches 454; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	MLHGASSRPATARKSSGLSGTVIRIPGDKSIHSRSMFGGLASGETRITGLLEGEDVINTG	60
DB	1	mshgassrpataarkssglsgrtviripgdkshsrsmfmgglasgetritgllegedvintg	60
QY	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDFGNAATGCRLTGVLGVYDFDSTFI	120
DB	61	kamqamgarirkegdtwiidgvnggllapeapldfngaatgcrltmglvgydfdstfi	120
QY	121	GDASLTRKPMGRVLNPLREMVGQVKSEGDRLPVTLRGPKTPPTITYRVPMASQVKSAN	180
DB	121	gdasltrkpmgrvlnplremvgvqksedgdrplvtlrgpktpptityrvpmasaqvksav	180
QY	181	LLAGLNTPGITVTIEPTWTRDHTKMLQGFAGNLTVETDADGVRTIRLEGRKLTGQVID	240
DB	181	llaglnpvgitvtiepimtrdhtekmlqgfganltvetdadvrtirlegrkltgqvld	240
QY	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPRTTGILILQEMGADIEVINPRLAGGED	300
DB	241	vpgdpsstafplvaallvpgsdvtilnvlmptrtgililtqemgadievlnprlagged	300
QY	301	VADLRVRSSTLKGVTVPEDRAPSMIDEPILAVAAFAEGATVMNGLEELRVKESDRLSA	360
DB	301	vadlrvrstlkgvtvpedrapsmidpilaavaafaegatvmngleelrvkesdrlsa	360
QY	361	VANGKLNGVDCDEGETSLVVRGPDGKGLGNASGAAVATHLDRHRIAMSFVLVGLVSEN	420
DB	361	vangklngvdcdegetslvvrpgdgkglgnasgaavathldhriamsflvmglvsenp	420
QY	421	VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA	455
DB	421	vtvdatmiatsfpefmdlmaglakieisdtkaa	455
RESULT 6			
ID	AAW24474		
XX	AAW24474 standard; Protein; 455 AA.		
AC	AAW24474;		
XX			
DT	02-OCT-1997 (first entry)		
XX			



Thu Aug 15 13:49:55 2002

1	msghasrpatarkssglsgtvripgdkksishrsfmfglasgetritgllegeadvintg	60
61	kamqamgarlrkkgdrtwiidgvcngllapeapldfpgnaatcrltmglvgyvdfdsfifi	120
61	kamqamgarlrkkgdrtwiidgvcngllapeapldfpgnaatcrltmglvgyvdfdsfifi	120
121	gdasltkrpgrvlnplremgvoqvksedgdrlpvtlurgpktpptirypmasaovksav	180
121	gdasltkrpgrvlnplremgvoqvksedgdrlpvtlurgpktpptirypmasaovksav	180
181	llaglnptpgittviepimtrdhtekmlqgfganltvettdadgvrtrlegrgkltgqvpid	240
181	llaglnptpgittviepimtrdhtekmlqgfganltvettdadgvrtrlegrgkltgqvpid	240
241	vpgdpstafplvaallvpgsdvtilnvlmnprrtclgltlqbwmgadiievnpriaggdb	300
241	vpgdpstafplvaallvpgsdvtilnvlmnprrtclgltlqbwmgadiievnpriaggdb	300
301	vadlrvrssltklgvtvpedrapsmidexpilavaaaafagatvmngleelrvkesdrlsa	360
301	vadlrvrssltklgvtvpedrapsmidexpilavaaaafagatvmngleelrvkesdrlsa	360
361	vanglklingvdcdegetslvyvrgrpdgkglgnasgaavathldhrtamsgflvmglvsenp	420
361	vanglklingvdcdegetslvyvrgrpdgkglgnasgaavathldhrtamsgflvmglvsenp	420
421	vtvddatmiatstfpfemdlmaglakielstdtkaa	455
421	vtvddatmiatstfpfemdlmaglakielstdtkaa	455

RESULT	8
AAE05053	
ID	AAE05053 standard; Protein; 455 AA.
XX	
AC	AAE05053;
XX	
DT	10-SEP-2001 (first entry)
XX	
DE	Agrobacterium sp. strain CP4 class II EPSPS protein.
XX	
XX	5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW	transformed bacteria; transgenic plant; herbicide.
KW	
XX	
OS	Agrobacterium sp. CP4.
XX	
PN	US6248876-B1.
XX	
PD	19-JUN-2001.
XX	
PF	20-AUG-1998; 98US-0137440.
XX	
XX	13-SEP-1994; 94US-0306063.
PR	07-APR-1997; 97US-0833485.
PR	31-AUG-1990; 90US-0576537.
PR	28-AUG-1991; 91US-0749611.
XX	
PA	(MONS ) MONSANTO CO.
XX	
PI	Barry GF, Kishore GM, Padgett SR, Stallings WC;
XX	
DR	WPI; 2001-407326/43.
DR	N-PSDB; AAD09754.
XX	
XX	DNA probe capable of use in a polymerase chain reaction for identifying
PT	the presence of a target genomic DNA encoding a
PT	5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme -
XX	
PS	Claim 3; Fig 3; 152pp; English.
XX	
CC	the present invention relates to a DNA probe capable of use in a
CC	polymerase chain reaction for identifying the presence of a target

CC	genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)					
CC	enzyme referred to as class II EPSPS enzyme which is tolerant to					
CC	glyphosate. EPSPS genes are useful in producing transformed bacteria					
CC	and transgenic plants which are tolerant to glyphosate herbicide. The					
CC	probe is useful for identifying the presence of a target genomic DNA					
CC	encoding a EPSPS enzyme. The present sequence is Agrobacterium sp.					
CC	strain CP4 class II EPSPS protein.					
CC	Note: The present sequence, SEQ ID NO: 3 is stated as amino acid					
CC	sequence throughout the specification. However, it is referred as probe					
CC	in claim 3 of the specification.					
XX						
SQ	Sequence	455 AA;				
<hr/>						
	Query Match	99.7%;	Score	2282;	DB	22; Length 455;
	Best Local Similarity	99.8%;	Pred. No.	1.le-180;		
	Matches	454;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
<hr/>						
QY	1	MLHGASSRPATARKSSGLSGTVIRIPGDKSIHSRSPMFGGLASGETRIITCLLEGEDVINTG	60			
DB	1					
QY	1	mshgassrpataarksgslsgtviripgdksishrsfmgfllasgetritcllegedvintg	60			
DB	61	kamqamgarirkegdwtiildvgngglilaPEAPLDFGNATCRLTMGLGVYDFDSTFI	120			
DB	61	kamqamgarirkegdwtiildvgngglilaPEAPLDFGNATCRLTMGLGVYDFDSTFI	120			
QY	121	GDA SLTKRPMGRVLNPLRMGMVOVKSEGDRLPVTLRGPKTPTITYRVPMA SAQVKS AV	180			
DB	121	gda sltkrpmgrvl nplrmgmvo vkse gdr l pvtl rlgpk tptit yrvpma saqvks av	180			
QY	181	LILAG LNTPGTITVIEPIIMTRDHTEKMLQGFGANLT VETDADGVRTIRLEGRKL TGQVID	240			
DB	181	lilag lntpgtittviepiimtrdh tekmlqfganltvetdadvrtirl egrkl tgqv id	240			
QY	241	VPGDSPSTAFPLVAALLVPGSDVTLNLVNMPRTGLI LTLOEMGADI EWINPRLAGGED	300			
DB	241	vpgdspstafplvaall vpgsdv tlnlvnmp rtgli ltloemga die winprlag ged	300			
QY	301	VADLRVRSSLTKGVTVPEDRAPSMIDEXYPILA VAAFAFAGATVMNGLEELRVKESDRLSA	360			
DB	301	vadlrvrssl tkgtvp edrap smidexy pilavaa fafaga tv mngleelrvkes drlsa	360			
QY	361	VANGLKLVDCDEGETSLVVRGPDGKGLGNASGA AVATHLDHRIAMSFVLMGLVSEN P	420			
DB	361	van glklingvdcde gets lvvrpgdg klglnas gaavathld hriams flvmglvsenp	420			
QY	421	VTVD DATMTATSPP EFMDLMAGLGI ELSDTKAA	455			
DB	421	vtvd datmtats pp efmdlmagl gi elsdtkaa	455			
<hr/>						
RESULT	9					
AAR22301	ID	AAR22301 standard; Protein; 449 AA.				
XX	AC	AAR22301;				
XX	DT	03-AUG-1992 (first entry)				
XX	DE	Class II EPSPS enzyme.				
XX	KW	Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;				
XX	KW	5'-enolpyruvylshikimate-3-phosphate synthase; weed control.				
XX	OS	Achromobacter sp. strain LBAA.				
XX	PN	WO9204449-A.				
XX	PD	19-MAR-1992.				
XX	PF	28-AUG-1991; 9LWO-US06148.				
XX	PR	31-AUG-1990; 9OUS-US076537.				

XX PA (MONS ) MONSANTO CO.  
XX PI Barry GF, Kishore GM, Padgett SR;  
XX PD WPI; 1992-114356/14.  
XX PF DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate  
XX PT synthase - for producing plants and bacteria tolerant to  
XX PT glyphosate herbicides  
XX PS Disclosure; Fig 5; 148pp; English.  
XX CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3  
XX CC phosphate synthase enzyme (EPSPS) of Achromobacter sp. strain LBAA  
XX CC It is used to create glyphosate resistant plants or seeds which  
XX CC can be planted in a field of crops to selectively control weeds.  
XX CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,  
XX CC tobacco and alfalfa. This provides a cost effective, environmentally  
XX CC compatible weed control device. See also AAR22300 and AAR22302.  
XX SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;  
Best Local Similarity 82.9%; Pred. No. 4.3e-149;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRPATARKSGLSGTVIRPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60  
DB 1 mshasapktararseatlgeiripgdkshrsfmgglasgetritgllegedvintg 60  
QY 61 KAMQAMGARIRKEGDWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120  
DB 61 ramqangakirkegdwiiingvngcllqpeaaldfgnagtgarlmglvgtymktsfi 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGRLPVTLRGPKTPPTITRVPMASAOVKSAV 180  
DB 121 gdaslkrpnmgrvnlplremgvqveaadgrmpitligpktanpityrvpmasaavksav 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVRTIRLEGRGLTGQVID 240  
DB 181 llaglnptgvtvteipvmtrdhtekmlqfgadltvetdkdgvhririgtggklvgqtlid 240  
QY 241 VPDPSSTAPPLVAALLVPGSDVITLNLNMPRTGLTLQEMGADIEVINPRLAGGED 300  
DB 241 vpdpsstapplvaallvpgsdvitrnlmptrtglitlqemgadievlnarlagged 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360  
DB 301 vadlrvrasklkgvvpvperapasmideypvlaaasfaegetvmdgldelrvkesdrlaa 360  
QY 361 VANGLKLVGDCDEGETSLVVRGPRPGKGLGNASGAATAVTHLDRHTAMSLVNGLVSENP 420  
DB 361 vangleangvdcdegetslvvrgrprpgkglgnasgaavathldhrtamsflvnglvsenp 420  
QY 421 VTVDATMIATSPFEPMDLMAGLAKIELS 450  
DB 418 vtvdnsmiatsfepfmdmmpglgakiels 447

RESULT 10  
AAR22302  
ID AAR22302 standard; Protein; 449 AA.  
XX  
XX AC AAR22302;  
XX DT 03-AUG-1992 (first entry)  
XX DE Class II EPSPS enzyme.  
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;  
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

XX OS Pseudomonas sp. strain PG2982.  
XX PN WO9204449-A.  
XX PD 19-MAR-1992.  
XX PF 28-AUG-1991; 91WO-US06148.  
XX PR 31-AUG-1990; 90US-0576537.  
XX PA (MONS ) MONSANTO CO.  
XX PI Barry GF, Kishore GM, Padgett SR;  
XX PD WPI; 1992-114356/14.  
XX CC DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate  
XX CC synthase - for producing plants and bacteria tolerant to  
XX CC glyphosate herbicides  
XX PS Disclosure; Fig 7; 148pp; English.  
XX CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3  
XX CC phosphate synthase enzyme (EPSPS) of Pseudomonas sp. strain PG2982  
XX CC It is used to create glyphosate resistant plants or seeds which  
XX CC can be planted in a field of crops to selectively control weeds.  
XX CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,  
XX CC tobacco and alfalfa. This provides a cost effective, environmentally  
XX CC compatible weed control device. See also AAR22300 and AAR22301.  
XX SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;  
Best Local Similarity 82.9%; Pred. No. 4.3e-149;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY 1 MLHGASRPATARKSGLSGTVIRPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60  
DB 1 mshasapktararseatlgeiripgdkshrsfmgglasgetritgllegedvintg 60  
QY 61 KAMQAMGARIRKEGDWIIDGVNGGLLAPLDFGNATGRLTMGLVGVYDFDSTFI 120  
DB 61 ramqangakirkegdwiiingvngcllqpeaaldfgnagtgarlmglvgtymktsfi 120  
QY 121 GDASLTKRPMGRVNLPLREMGVQVSKEDGRLPVTLRGPKTPPTITRVPMASAOVKSAV 180  
DB 121 gdaslkrpnmgrvnlplremgvqveaadgrmpitligpktanpityrvpmasaavksav 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQFGANLTVETDADGVRTIRLEGRGLTGQVID 240  
DB 181 llaglnptgvtvteipvmtrdhtekmlqfgadltvetdkdgvhririgtggklvgqtlid 240  
QY 241 VPDPSSTAPPLVAALLVPGSDVITLNLNMPRTGLTLQEMGADIEVINPRLAGGED 300  
DB 241 vpdpsstapplvaallvpgsdvitrnlmptrtglitlqemgadievlnarlagged 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYPLAVAAAFAGCATVMNGLEELRVKESDRLSA 360  
DB 301 vadlrvrasklkgvvpvperapasmideypvlaaasfaegetvmdgldelrvkesdrlaa 360  
QY 361 VANGLKLVGDCDEGETSLVVRGPRPGKGLGNASGAATAVTHLDRHTAMSLVNGLVSENP 420  
DB 361 vangleangvdcdegetslvvrgrprpgkglgnasgaavathldhrtamsflvnglvsenp 420  
QY 421 VTVDATMIATSPFEPMDLMAGLAKIELS 450  
DB 418 vtvdnsmiatsfepfmdmmpglgakiels 447

RESULT 11  
AAW34684

AAW34684 standard; Protein; 449 AA.

AAW34684;

17-FEB-1998 (first entry)

Class II EPSP synthase (EPSPS) from *Achromobacter* sp. strain LBAA.

5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate; shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid; 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant; glyphosate resistance gene; glyphosate-tolerance; promoter.

*Achromobacter* sp. strain LBAA.

Key	Location/Qualifiers
Region	200..204
FT	/label= characteristic_region
FT	/note= "see AAW34690"
Region	26..29
FT	/label= characteristic_region
FT	/note= "see AAW34691"
Region	173..177
FT	/label= characteristic_region
FT	/note= "see AAW34692"
Region	271..274
FT	/label= characteristic_region
FT	/note= "see AAW34693"
XX	
XX	
PN	US5627061-A.
PD	06-MAY-1997.
XX	
XX	
XX	07-JUN-1995; 95US-0476008.
XX	
PR	13-SEP-1994; 94US-0306063.
PR	31-AUG-1990; 90US-0576537.
PR	28-AUG-1991; 91US-0749611.
XX	
XX	(MONS ) MONSANTO CO.
XX	
XX	Barry GF, Kishore GM, Padgett SR, Stallings WC;
PI	WPI; 1997-271315/24.
DR	N-PSDB; AAT93789.
XX	
PT	Production of glyphosate-herbicide tolerant plants - using DNA
PT	encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT	enzyme(s)
XX	
XX	Disclosure; Fig 4; 151pp; English.
XX	
CC	AAW34683-89 represent a new class of glyphosate-tolerant
CC	5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC	EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC	belong to a new class, Class II. The present sequence was isolated from
CC	<i>Achromobacter</i> sp. strain LBAA. The EPSPS enzymes are part of the
CC	shikimic acid pathway, which leads to the biosynthesis of aromatic
CC	compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and
CC	3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC	inhibited by the herbicide glyphosate. It would be useful to produce
CC	transgenic crops containing glyphosate resistance genes so that
CC	glyphosate-containing herbicides can be applied to selectively kill
CC	weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for
CC	glyphosate, such that when introduced into a plant, the plant is made
CC	glyphosate-tolerant, and EPSPS enzyme activity is not affected. These
CC	glyphosate-tolerant, and EPSPS enzymes are fused to a chloroplast transit peptide to
CC	Class II EPSPS enzymes are fused to a chloroplast transit peptide to
CC	target the protein into the chloroplast, which is the site for the
CC	shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant
CC	under the control of a promoter such as figwort mosaic virus promoter or
CC	the cauliflower mosaic virus promoter, so that expression is enhanced.
XX	
XX	449 AA.

Query Match 83.1%; Score 1900.5; DB 18; Length 449;

Best Local Similarity 82.9%; Pred. No. 4.3e-149;

Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

QY	1	MLHGASRRPATARKSGSLGVTRIFGDKSISHRSMFGLASGETRITGLLEGEDVINTG	60
Db	1	mshsaspkpatarrsealtgeiripgdkshsrsmfmgglasgetritgllegedvintg	60
QY	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPLDFGNAATGCRUTMGLGVYDFDSTFI	120
Db	61	ramqamgakirkegdwiingvngccllqpeaalofgnagcgarltmglvgtymktsfi	120
QY	121	GDASLTKRPMGRVLPNPLREMGVQVKSESDGRLPVTLRGPKTPPTTYRVPMSAQVKSASV	180
Db	121	gdaslskrpmgrvlnplremgvqveaadgrmpltligktpanpttyrvpmasaqvksav	180
QY	181	LLAGLNPGTITVIEPIMTRDHTKMLQGFANLIVETDADGVTRIRLEGRGKLTGVQVID	240
Db	181	llaglnptgtvievpmtrdhtkmlqgfadltvetdkdgkvlrirtgqgklvgqgvid	240
QY	241	VPGPSSTAFPLVAALLVPSDVTILNVLNPTRTGLILTQEMGADIEVINPRLAGGED	300
Db	241	vpgpsstaflpvaallvpsdvtlrvnlmpttrtgliltqemgadievlnarlagged	300
QY	301	VADLRVRSSTLKGTVTPEDRAFSPMIDEXYPILAVAAAFAGATVMNGLELRVKESDRLSA	360
Db	301	vadlrvrssltkgtvtpedrafspmidexypilavaaafagatvmnglelrvkessdrlsa	360
QY	361	VANGLKLVDCDEGETSLVVRGRPDGKGNASGAATAATHLDHRTAMSLFVGLVSNRP	420
Db	361	vargleangvdcdegetslvvrgrpdgkgnasgaataathldhrtamslfvlvsnrp	420
QY	421	VTVDDATMIATSPPEFMDLMAGLAKIELS	450
Db	418	vtvddsnmiatspfevdmmmpglgakielS	447

RESULT 12

AAW34685

ID AAW34685 standard; Protein; 449 AA.

XX AAW34685;

XX AAW34685;

DT 17-FEB-1998 (first entry)

XX Class II EPSP synthase (EPSPS) from *Pseudomonas* sp. strain PG2982.

XX 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;

XX shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;

XX 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;

XX glyphosate resistance gene; glyphosate-tolerance; promoter.

XX *Pseudomonas* sp. strain PG2982.

Key	Location/Qualifiers
Region	200..204
FT	/label= characteristic_region
FT	/note= "see AAW34690"
Region	26..29
FT	/label= characteristic_region
FT	/note= "see AAW34691"
Region	173..177
FT	/label= characteristic_region
FT	/note= "see AAW34692"
Region	273..276
FT	/label= characteristic_region
FT	/note= "see AAW34693"
XX	
XX	US5627061-A.
XX	06-MAY-1997.



Query Match 83.1%; Score 1900.5; DB 18; Length 449;  
Best Local Similarity 82.9%; Pred. NO. 4.3e-149;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

Qy 181 LLAGINTPGITTVIEPIETMRDHTKMLQGGFANLTVDADGVRTIRLEGRGKLTGOVID 240  
Db 181 lllagintpgvttvlepvmttrdhtekmlqggfadtvetdkdgvhrhritgqgklvggtid 240

Qy 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPRTTGLLTLOEMGADIEVINPRLAGGED 300  
Db 241 vpgdpsstafplvaallvegsvdvtirnlvmptrtgllltlqemgadlevlnarlagged 300

Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEXPIILAAFAAFAEGATVMNGLEELRVKESDRLSA 360  
Db 301 vadlrvrasklkgvvppperapsmideypvllaasfaegetvmdgidelrvkesdrila 360

Qy 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVTHLDRHIAFSFLVGLVSEN 420  
Db 361 vangleangvdcetegemsltvrgpdgkglg---ggvtathldhriamsflvmglaaekp 417

Qy 421 VTVDATMIATSFPEFMDLMAGLGAKIELS 450  
Db 418 vtvdnsnmatsfpefmdmpgglgakiels 447

RESULT 14  
AAW24480  
ID AAW24480 standard; Protein; 449 AA.  
XX AC AAW24480;  
XX DT 02-OCT-1997 (first entry)  
XX DE Class II EPSPS used for glyphosate resistant plant production.  
XX KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;  
XX KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;  
XX KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;  
XX KW tobacco.  
XX OS Pseudomonas sp. strain PG2982.  
XX PN US5633435-A.  
XX PD 27-MAY-1997.  
XX PF 31-AUG-1990; 90US-0576537.  
XX PR 13-SEP-1994; 94US-0306063.  
XX PR 31-AUG-1990; 90US-0576537.  
XX PR 28-AUG-1991; 91US-0749611.  
XX PA (MONS ) MONSANTO CO.  
XX PI Barry GF, Kishore GM, Padgett SR, Stallings WC;  
XX DR WPI; 1997-297418/27.  
XX DR N-PSDB; AAW77316.  
XX PT New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene  
XX PT used for transforming plants to produce plants which are tolerant to  
XX PT glyphosate herbicide  
XX PS Claim 7; Column 69-72; 154pp; English.  
XX CC AAW24480 shows the sequence of a class II 5-enolpyruvylshikimate-3-  
XX CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant  
XX CC to glyphosate herbicides. EPSPS and sequences encoding it are used for  
XX CC the production of herbicide resistant (glyphosate-tolerant) plants  
XX CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed  
XX CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,  
XX CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.  
XX SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 18; Length 449;  
Best Local Similarity 82.9%; Pred. NO. 4.3e-149;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

Qy 1 MLHGASSRPATARKSGSLGSGTIVRIPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60  
Db 1 mshsaaspkpatarrsealtgeiripgdkshrsfmgfllasgetritgllegedvintg 60

Qy 61 KAMQAMCARIKKEDTWIIDGVNGGLLAPLDFGNAATGCLTMGLVGVDFDSTFI 120  
Db 61 ramqamgakarkegdvwlvgngcillqpeaaldfgnagtgarltnglvqtdmktfsi 120

Qy 121 GDASLTRKPRGRVLNPLREMGVQVKSEGDRLPVLTRGPKTPTPTITVYRVPMAASQVKS 180  
Db 121 gdaalskrpgrvlnplremgvqveaadgrmpitllgpktpitvyrvpmasaqqksav 180

Qy 181 LLAGINTPGITTVIEPIETMRDHTKMLQGGFANLTVDADGVRTIRLEGRGKLTGOVID 240  
Db 181 lllagintpgvttvlepvmttrdhtekmlqggfadtvetdkdgvhrhritgqgklvggtid 240

Qy 241 VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPRTTGLLTLOEMGADIEVINPRLAGGED 300  
Db 241 vpgdpsstafplvaallvegsvdvtirnlvmptrtgllltlqemgadlevlnarlagged 300

Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEXPIILAAFAAFAEGATVMNGLEELRVKESDRLSA 360  
Db 301 vadlrvrasklkgvvppperapsmideypvllaasfaegetvmdgidelrvkesdrila 360

Qy 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATAVTHLDRHIAFSFLVGLVSEN 420  
Db 361 vangleangvdcetegemsltvrgpdgkglg---ggvtathldhriamsflvmglaaekp 417

Qy 421 VTVDATMIATSFPEFMDLMAGLGAKIELS 450  
Db 418 vtvdnsnmatsfpefmdmpgglgakiels 447

RESULT 15  
AAW71611  
ID AAW71611 standard; Protein; 449 AA.  
XX AC AAW71611;  
XX DT 14-DEC-1998 (first entry)  
XX DE Pseudomonas sp. strain PG2982 Class II EPSPS.  
XX KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;  
XX KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;  
XX KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;  
XX KW N-phosphonomethylglycine.  
XX OS Pseudomonas sp.  
XX PN US5804425-A.  
XX PD 08-SEP-1998.  
XX PF 07-APR-1997; 97US-0833485.  
XX PR 13-SEP-1994; 94US-0306063.  
XX PR 31-AUG-1990; 90US-0576537.  
XX PR 28-AUG-1991; 91US-0749611.  
XX PR 07-APR-1997; 97US-0833485.  
XX PA (MONS ) MONSANTO CO.  
XX PI Barry GF, Kishore GM, Padgett SR, Stallings WC;  
XX DR WPI; 1998-505657/43.  
XX DR N-PSDB; AAW58012.  
XX PT Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -



PT useful for characterisation of the enzyme to determine inhibition  
PT data values  
XX  
PS Disclosure; Fig 5; 152pp; English.  
XX  
CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the  
CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),  
CC -G-D-K-X3- (II); -S-A-O-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,  
CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,  
CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,  
CC produced by recombinant methods, can be used in kinetic studies to  
CC determine Ki and Km values of the enzyme for its characterisation. The  
CC enzyme is normally used for the production of 5-enolpyruvyl-3-  
CC phosphoshikimic acid in plants, and most forms of the enzyme are  
CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.  
CC Inhibition data enables more accurate values of concentrations of  
CC herbicide to be used when growing the plant without being detrimental to  
CC it. This enables the plant to be grown in the presence of the herbicide,  
CC being used to inhibit the growth of undesired plants. The present  
CC sequence represents a Class II EPSPS from bacterial isolate  
CC Pseudomonas sp. strain PG2982.  
XX  
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 19; Length 449;  
Best Local Similarity 82.9%; Pred. No. 4.3e-149;  
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

Qy 1 MLHGASSRPATARKSSGLSGTVIRPGDKSISHRSFPMFGLASGETRITGLLEGEDVINTG 60  
Db 1 mshaspkpatarrsealtgeiripgdkshrsfmgfglasgetritgllgedvintg 60

Qy 61 KAMQAMGARIRKEGDTIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120  
Db 61 ramqamgakirkedvlingvngccllqpeaaldfnagtgartlmglvgtymktsfi 120

Qy 121 GDASLTRPMGRVNLPLREMGVQVKSEGDRLPTVTLRGPKTPTITYRVPWASQVKSAY 180  
Db 121 gdaalskrpmsgrvnlplremgvqveaadgdrmpiltligpkcanpityrvpmasaqvksav 180

Qy 181 LLAGINTPGITVTIEPIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRKLTGQVID 240  
Db 181 llagintpgvtviepvmtrdhtekmlqgfgaadltvetdkdgvhrhritggklyvgqtid 240

Qy 241 VPGDPSSTAFPLVAALLYPGSDVTILNVMNPTRTGLTLTQEMGADIEVINPRLAGGED 300  
Db 241 vpgdpsstafplvaallyvegsvdvtirnlmnptrtgltiltqemgadievlnarlagged 300

Qy 301 VADLRVRSSTLKGVTVPEDRAPSMTIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
Db 301 vadlrvrasklkgvvppperapsmideypvliiaasfaegetvmgdgldeilrvkesdrila 360

Qy 361 VANGKLKNGVDCDEGETSLVYVRGPRDGKGLGNASGAAVATHLDHRIAMSFVLMGLVSENP 420  
Db 361 vangleangvdcetegemsltvrgprdgkglg---ggvtathldhriamsflvmglvseakp 417

Qy 421 VTVDATWIAATSFPEFMDLMAGLAKIELS 450  
Db 418 vtvdnsnmiatsfpefmdmpglgakiels 447

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:21 ; Search time 17.59 Seconds  
(without alignments)  
1001.557 Million cell updates/sec

Title: US-09-464-099A-3

Perfect score: 2288

Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMACLGAKIELSDTKAA 455

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	1 AROA_AGRSP	Q9r4e4 agrobacteri
2	1906.5	83.3	449	1 AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1 AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1 AROA_SYNV3	Q59975 synecocyst
5	806.5	35.2	430	1 AROA_LACLA	Q9ce00 lactococcus
6	806	35.2	443	1 AROA_BACNO	Q94550 bacteroides
7	803.5	35.1	427	1 AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1 AROA_LACLC	P20691 bacillus su
9	785	34.3	428	1 AROA_BACSU	Q67494 aquifex ae
10	766.5	33.5	431	1 AROA_AQUAE	Q9wy10 thermotoga
11	695	30.4	410	1 AROA_THEMEA	P52312 campylobact
12	590	25.8	428	1 AROA_CAMJE	Q05615 staphylococ
13	577	25.2	430	1 AROA_STAAD	Q92kf7 helicobacte
14	521.5	22.8	429	1 AROA_HELPJ	P56197 helicobacte
15	518.5	22.7	429	1 AROA_HELPY	O26860 methanobact
16	496	21.7	419	1 AROA_METTH	Q57925 methanococc
17	459.5	20.1	429	1 AROA_METJA	Q9rvd3 deinococcus
18	408	17.8	439	1 AROA_DEIRA	O28775 archaeoglob
19	403	17.6	416	1 AROA_ARCFU	Q9yek9 aeropyrum p
20	402.5	17.6	427	1 AROA_AERPE	P19688 yersinia en
21	402.5	17.6	427	1 AROA_YEREN	O60112 yersinia pe
22	395	17.3	428	1 AROA_VERPE	P52310 haemophilus
23	387.5	16.9	432	1 AROA_HAESO	P07638 escherichia
24	378	16.5	427	1 AROA_ECOLI	Q9zff7 shigella so
25	376	16.4	427	1 AROA_SHISO	Q9v1h1 pyrococcus
26	373.5	16.3	410	1 AROA_PYRAB	P54220 pasteurella
27	373	16.3	432	1 AROA_PASHA	P24497 klebsella
28	372	16.3	427	1 AROA_KLEPN	O04570 pasteurella
29	372	16.3	440	1 AROA_PASMD	O87006 shigella dy
30	371	16.2	427	1 AROA_SHIDI	P39915 burkholderi
31	365.5	16.0	463	1 AROA_BURPS	Q9x4h2 edwardsiell
32	365	16.0	428	1 AROA_EDWIC	Q03421 haemophilus
33	363.5	15.9	432	1 AROA_HAEIN	

34 357 15.6 427 1 AROA\_SALTY P07637 salmonella  
35 348 15.2 427 1 AROA\_SALGL P22299 salmonella  
36 347 15.2 427 1 AROA\_SALTI P19786 salmonella  
37 323.5 14.1 516 1 AROA\_PETHY P11043 petunia hyb  
38 321.5 14.1 1588 1 ARO1\_YEAST P08566 s pentafunc  
39 317 13.9 441 1 AROA\_CHLMU Q9pk28 chlamydia m  
40 315.5 13.8 518 1 ARO1\_TOBAC P23981 nicotiana t  
41 314.5 13.7 440 1 AROA\_CHLTR O84371 chlamydia t  
42 313 13.7 428 1 AROA\_BUCAP Q59178 buchnera ap  
43 310.5 13.6 520 1 AROA\_LYCES P10748 lycopersico  
44 308.5 13.5 1573 1 ARO1\_SCHPO Q9p7r0 s pentafunc  
45 306.5 13.4 445 1 AROA\_CHLPN Q9z6m0 chlamydia p

#### ALIGNMENTS

RESULT 1

ARO\_AGRSP STANDARD; PRT; 455 AA.

AC Q9r4e4;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-

DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).

GN AROA.

OS Agrobacterium sp. (strain Cp4).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Agrobacterium.

OX NCBI\_TaxID=361;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.

RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;

RT "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";

RL Patent number US5633435, 27-MAY-1997.

RN [2]

RP SEQUENCE OF 1-15.

RX MEDLINE=96182485; PubMed=8598558;

RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,

Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,

Fuchs R.L., Padgett S.R.;

"The expressed protein in glyphosate-tolerant soybean, 5-

enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp.

strain Cp4, is rapidly digested in vitro and is not toxic to acutely

gavaged mice.";

RL J. Nutr. 126:728-740(1996).

CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =

phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.

CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN

THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -!- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in

glyphosate-tolerant soybean, canola, cotton and maize by Monsanto.

Developed to provide new weed-control options for farmers.

Expression of this protein in plants imparts high levels of

glyphosate tolerance.

CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

DR Interpro: IPR001986; EPSP\_synthase.

DR Pfam: PF00275; EPSP\_synthase; 1.

DR PROSITE: PD001867; EPSP\_synthase; 1.

DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.

DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; 1.

KW Aromatic amino acid biosynthesis; Transferrase; Herbicide resistance;

Genetically modified food.

FT CONFLICT 2 S -> L (IN REF. 2).

SQ SEQUENCE 455 AA; 47588 MW; 236580D08DE422 CRC64;

Query Match

Best Local Similarity 100.0%; Score 2288; DB 1; Length 455;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSHGASSRPATARKSSGLSGTVPKSGTSHRSFMPFGLASGETRITGLLEGEDVINTG 60
DB 1 MSHGASSRPATARKSSGLSGTVPKSGTSHRSFMPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCLRTMGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCLRTMGLVGVYDFDSTFI 120
QY 121 GDASLTRKPRMGRVNLNPLREMGVQVKSSEDGRLPVTLRGPKTPTPTTYRVPMAQAQVSAV 180
DB 121 GDASLTRKPRMGRVNLNPLREMGVQVKSSEDGRLPVTLRGPKTPTPTTYRVPMAQAQVSAV 180
QY 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGANLTVETDADGVRTIRLEGRGKLTGGVID 240
DB 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGANLTVETDADGVRTIRLEGRGKLTGGVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLHRIAMSFVLMGLVSEN 420
DB 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLHRIAMSFVLMGLVSEN 420
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA 455
DB 421 VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA 455

```

Query Match 83.3%; Score 1906.5; DB 1; Length 449;  
 Best-Local Similarity 83.1%; Pred. No. 1.8e-111;  
 Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```

QY 1 MSHGASSRPATARKSSGLSGTVPKSGTSHRSFMPFGLASGETRITGLLEGEDVINTG 60
DB 1 MSHSASPATARRSEALTGEIRIPGDKSISHRSFMPFGLASGETRITGLLEGEDVINTG 60
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCLRTMGLVGVYDFDSTFI 120
DB 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFGNAATGCLRTMGLVGVYDFDSTFI 120
QY 121 GDASLTRKPRMGRVNLNPLREMGVQVKSSEDGRLPVTLRGPKTPTPTTYRVPMAQAQVSAV 180
DB 121 GDASLTRKPRMGRVNLNPLREMGVQVKSSEDGRLPVTLRGPKTPTPTTYRVPMAQAQVSAV 180
QY 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGANLTVETDADGVRTIRLEGRGKLTGGVID 240
DB 181 LLAGLNTPGITTVTIEPIMTRDHTKMLQGGANLTVETDADGVRTIRLEGRGKLTGGVID 240
QY 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300
DB 241 VPGDPSSTAFPLVAALLVPGSDVTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
DB 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAAFAGATVNGLEELRVKESDRLSA 360
QY 361 VANGKLNGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLHRIAMSFVLMGLVSEN 420
DB 361 VARGLEANGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLHRIAMSFVLMGLVSEN 420
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450
DB 421 VTVDATMIATSPFPEFMDLMAGLAKIELS 450

```

RESULT 3  
 AROA\_PSES2  
 ID AROA\_PSES2 STANDARD; PRT; 431 AA.  
 AC O9KCA6;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
 enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
 GN AROA OR AROE OR BHL667.  
 OS Bacillus halodurans  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
 phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
 CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE  
 BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AP001512; BAB05386.1; -  
 DR InterPro: IPR001986; EFSP\_syntase.  
 DR Pfam; PF00275; EFSP\_syntase; 1.  
 DR ProDom; PD001867; EFSP\_syntase; 1.  
 DR PROSITE; PS00104; EFSP\_SYNTHASE\_1; 1.  
 DR PROSITE; PS00885; EFSP\_SYNTHASE\_2; 1.  
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
 SQ SEQUENCE 431 AA; 45485 MW; 12F4FFBE7BA0743D CRC64;

Query Match 38.6%; Score 884; DB 1; Length 431;  
 Best Local Similarity 44.2%; Pred. No. 4.7e-48;  
 Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;

QY 15 SSGLSGTVRIPOGDKSISHRSPWFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRK 74  
 DB 10 AKGLGTTIKVPGDKSISHRVNFVGLAKGTTTVEGLFPGADCLSPISCFQKLGVSIEQAE 69  
 75 DTWIIDGVNGGGLLAPEAPLDFGNAATGCRLTMTGLVGVYDFDSTFGDASLTKRPMGRVL 134  
 70 ERYTVKGGWGLREPSDILDVNSGTTTLLILGILSLPFLSHVLIIGDESIGKRPWKRV 129  
 QY 135 NPLREMGVQVKSSED-GDRLPVTLRGPKTPPTITYRVPVMSAQVSAVLLAGLNTPGITTV 193  
 DB 130 EPLKSMGAQIDGRDHGNTPLSIRGGQL-KGIDFHPSPVSAQMKSAILLAGLRAEGKTSV 188  
 QY 194 IEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSAPPLV 253  
 DB 189 TEPARTDHTERTMLAEFAGVNI-----EKDGL-TVSIEGQMLTGQHVVPVPGDISAAFFLV 243  
 QY 254 AALVPGSDVTLNVLNMPRTGLTLTLOEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313  
 DB 244 AGAMVPHSRITLVNGINPTFRAGILEVUKMGATUAMENRVQGGEPVADLTFTSVLQG 303  
 QY 314 VTPEDRAPSMIDEYPIILAVAAFAEGATVAMNGLEELRVKESDRLSAVANGKLKGVDCD 373  
 DB 304 VEIGGDIIPRLIDEIPILAVLTAQSGRTVKDAELAKVETNRDITVVSLETLKLGASTH 363  
 QY 374 EGETSLVVRGPRDGGKLGNSAGAAVATHDRIAMSLFVGLVNSPNVTVDDATMIATSF 433  
 DB 364 ATDDGMIEGTPPLKG-----GTVSSHGDHRIGMAMAIAALLAEKPTVTEGTEIAVSY 418  
 QY 434 PEFMD 438  
 DB 419 PSFFD 423

JUL 4  
 ID AROA\_SVNY3  
 AC Q59975; Q59974; STANDARD; PRT; 447 AA.  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
 enolpyruvylshikimate-3-phosphate synthase) (EPSPS).  
 GN AROA OR SLR0444.  
 OS Synecocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94299161; PubMed=8026753;  
 RA dalla Chiesa M., Mayes S.R., Maskell D.S., Nixon P.J., Barber J.;  
 RT "An *aroA* homologue from *Synechocystis* sp. PCC 6803.";  
 RL Gene 144:145-146(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hiroseawa M., Suglura M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
*Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
 entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 RN [3]  
 RP SEQUENCE OF 103-194 FROM N.A.  
 RX MEDLINE=93307506; PubMed=7686511;  
 RA Mayes S.R., dalla Chiesa M., Zhang Z., Barber J.;  
 RT "The genes *aroA* and *trnQ* are located upstream of *psbO* in the  
 chromosome of *Synechocystis* 6803.";  
 RL FEBS Lett. 325:255-261(1993).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
 -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
 THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75325; CAA53074.1; -  
 DR EMBL; D90914; BAA18477.1; -  
 DR EMBL; X72784; CAA51291.1; -  
 DR InterPro; IPR001986; EFSP\_syntase.  
 DR Pfam; PF00275; EFSP\_syntase; 1.  
 DR ProDom; PD001867; EFSP\_syntase; 1.  
 DR PROSITE; PS00104; EFSP\_SYNTHASE\_1; 1.  
 DR PROSITE; PS00885; EFSP\_SYNTHASE\_2; 1.  
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
 SQ SEQUENCE 447 AA; 47046 MW; 2F3C8AD26B5A7BCE CRC64;

Query Match 37.9%; Score 867.5; DB 1; Length 447;  
 Best Local Similarity 48.1%; Pred. No. 5.2e-47;  
 Matches 203; Conservative 62; Mismatches 144; Indels 13; Gaps 7;

QY 18 LSGTVRIPOGDKSISHRSPWFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRK-EGDT 76  
 DB 26 LTGLRVPGDKSISHRALMLGAITGETIIEGLLGEDPRSTAHCFRAMGAETSELNSEK 85  
 QY 77 WIIDGVNGGGLLAPEAPLDFGNAATGCRLTMTGLV-GVYDFDSTFGDASLTKRPMGRVLN 135  
 DB 86 IIVQGRGLGQLQEPSTVLDAGNSGTTMLMLGLLAGQKDCLTFTVGDDSLRRHPMSRVIQ 145  
 QY 136 PLREMGVQV-KSEGDRLPVTLRGPKTPPTITYRVPVMSAQVSAVLLAGLNTPGITTVI 194  
 DB 146 PLOQMGAKIWARSKNGKFAFLAVQSGQL-KPIHYHSPIASAQVSKSLLAGLTTEGDTTV 204  
 QY 195 EPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGKGLTGQVIDVPGDPSAPPLV 254  
 DB 205 EPALSRDHSERMLQAFGAKLTID--PVTHSVTVHGHAPHLTGQVVVPGDISAAFWLVA 261  
 QY 255 ALLVPGSDVTLNVLNMPRTGLTLTLOEMGADIEVINPRLAGGEDVADLRVRSSTLKG 314  
 DB 262 ASILPGSELLVNGVINPTFRAGILEVUKMGATUAMENRVQGGEPVADLTFTSVLQG 321  
 QY 315 TYPEDRAPSMIDEYPIILAVAAFAEGATVAMNGLEELRVKESDRLSAVANGKLKGVDCD 374  
 DB 322 TFGGEIIPRLIDEIPILAVAAFAEGTTRIEDAAELRVKESDRLSAIASELCKMAKYTE 381  
 QY 375 GETSLVVRGPRDGGKLGNSAGAAVATHDRIAMSLFVGLVNSPNVTVDDATMIATSF 434  
 DB 382 FDDGLEIQG---GSLP---QGAENVSLDTHRIAMALATAALGSGGQTIINRAEAAAISSVP 435  
 QY 435 EF 436

Thu Aug 15 13:49:53 2002

us-09-464-099a-3.rsp

```

Db 436 EF 437

RESULT 5
AROA_LACLA
ID AROA_LACLA STANDARD; PRT; 430 AA.
AC Q9CEU0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
DE AROA OR LLI744.
GN Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILI403;
RC MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Welzenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis ILI403."
RL Genome Res. 11:731-753(2001).
CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate
CC -|- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006404; AAK05842.1;
CC InterPro; IPR001986; EPSP_synthase.
CC Pfam; PF00275; EPSP_synthase; 1.
CC ProDom; PD001867; EPSP_synthase; 1.
CC PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
CC PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferase; Complete proteome.
CC SEQUENCE 430 AA; 45958 MW; CB216F07AA4EE799 CRC64;

Query Match 35.2%; Score 806.5; DB 1; Length 430;
Best Local Similarity 41.9%; Pred. No. 3e-43;
Matches 184; Conservative 77; Mismatches 157; Indels 21; Gaps 7;

QY 15 SSGLSGTVRIPGDKSISHRFSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
Db 7 SOGLKGLKVPYDGSISHRFSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 66
QY 75 DTWIDGVNGVGLLAPRAPIDFNGNATGCRLTWGLVGVYDFDSFTFGDASLTAKRPMGRVL 134
Db 67 EVITVHGKIGSELKAPEKALDMGNSGTSTRLLSGILAGLPFTFLFGDDSLSRKPRMDRVA 126
QY 135 NPLEMGVQVQKSE-DGDRLPVTLRGKPTPTITRVPMASQVKSALLAGLNTPGITTV 193
Db 127 TPLQMGAEITGQTDKVKLPMTINGSTHLKADIVLVPASQVKSIVFAALQAEGLTKV 186
QY 194 IEPTWTRDHTKMLQGFGANLTETDADGVRTIRLEGRKGLTGQVIDVDPGDPSPSTAPPLV 253
Db 187 VEKEKTRSHTEMLVQFGGELKVS-----KTILVPGQKLVGQKVVVPGVDISSAAFVLV 241

```

# RESULT 6

```

AROA_BACNO
ID AROA_BACNO STANDARD; PRT; 443 AA.
AC Q46550;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
DE AROA.
GN Bacteroides nodosus (Dichelobacter nodosus).
OS Bacteria; Proteobacteria; gamma subdivision; Dichelobacteriaceae;
OC Dichelobacter.
OX NCBI_TaxID=870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VCS1001;
RC MEDLINE=94320795; PubMed=8045432;
RA Alm R.A., Dalrymple B.P., Mattick J.S.;
RT "Sequencing and expression of the aroA gene from Dichelobacter
RT nodosus."
RL Gene 145:97-101(1994).
CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -|- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29339; CAA82544.1;
CC InterPro; IPR001986; EPSP_synthase.
CC Pfam; PF00275; EPSP_synthase; 1.
CC ProDom; PD001867; EPSP_synthase; 1.
CC PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
CC PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
CC Aromatic amino acid biosynthesis; Transferase.
CC SEQUENCE 443 AA; 47444 MW; BE2243277ADEFB5 CRC64;

```

```

Query Match 35.2%; Score 806; DB 1; Length 443;
Best Local Similarity 42.2%; Pred. No. 3.4e-43;
Matches 185; Conservative 74; Mismatches 169; Indels 10; Gaps 3;

QY 16 SGLSGTVRIPGDKSISHRFSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 75
Db 12 SALSGETITCGDKSMHRALLLAALAEAGTEIRGLACADCLATQALRALGVDIQREKE 71

```







```

Db 239 AGAMVPSRIVLKNVGLNPTGIGTIDVQNMGAKEIKPSADSGAEPYGDLLIETSSLK 298
QY 313 GVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKNGVDC 372
Db 299 AVEIGGDIIPRLIDEPIIALLATQAEETGVTKDAAELVKETNRDITVVVSELKLGAEI 358
QY 373 DEGETSLVVRGPDGKGLGNASGAATVATHLDRHIAFLVGLVSENPTVDDATMIATS 432
Db 359 EPTADGMKYVGKOTLKG-----GAAVSSGHDHRIGMMLGFIASCITEPIEHTDAIHS 413
QY 433 FPEFMDMAGLGAK 446
Db 414 YTFPEHNLKLSK 427

RESULT 10
AROA_AQUAE STANDARD; PRT; 431 AA.
O67494;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR AQ.1536.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus";
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000744; AAC07443.1; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE.1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE.2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 431 AA; 47793 MW; EF842512EBE41D2A CRC64;

Query Match 33.5%; Score 766.5; DB 1; Length 431;
Best Local Similarity 40.7%; Pred. No. 9.2e-41;
Matches 174; Conservative 73; Mismatches 169; Indels 11; Gaps 6;

QY 14 KSSGLSGTVRPGDKSISRSRSMFGGLASGETRITGLLEGEDVINTGKAMQANGAIRKE 73
Db 6 KIKRVKGLRVPDSKSIITHRAFTLGAASGETLVKRPGLISGDTLATLEILKAIRTKVREG 65

```

```

QY 74 GDTWIIDGVNGGILLAPEDFGNATGCRLTWGLVGVYDFDSTFTIGDASLTFRPMGRV 133
Db 66 KEVLLTEG-RNYTFLEPHDVLDAKNSGTTARIMSGVLSTQPFPSVLTGDSLNRPMLRV 124
QY 134 LNPRLRMGVQVKS-EDGDRLPVTLRGPKTPTPTIYRVPMAAQVKSASVILAGLNTPGITT 192
Db 125 VEPRLMAGKIDREBGNKLPPIAIRGNL-KGISYFNKKSAAQVKSALLLAGLRAEGMTE 183
QY 193 VIEPIMTRDHTKMLQGFGANLTAVETDADGVTRIRLEGRGKLTGOVLDVDPGDSSTAFPL 252
Db 184 VVEPYLSRDHTERMLKLFAGAEVITPEERG-HIVKIKGGQELQGTVEYCPADPSSAAFYA 242
QY 253 VAALLVPGSDVTILNVLNPNTRTGLIILTQEMGADIEVINPRLAGGEDVADLRVRS-STL 311
Db 243 ALATLAPEGEIRLKEVLLNPNTRDGFYRKLIEMGGDISFENYRELSNPNMADLVVRPVDNL 302
QY 312 KGVTPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKNGVD 371
Db 303 KPVKVSPEEVPITLIDEIPILAVLMAFADGVSEVKAGELVKESDRIKAIVTNLRKLGQV 362
QY 372 CDEGETSLVVRGPDGKGLGNASGAATVATHLDRHIAFLVGLVSENPTVDDATMIAT 431
Db 363 VEEFEDGFATHGTRKEIK-----GGVITFKDHRIMAFVGLVVEEVIIDHPECVTY 416
QY 432 SFPEFMD 438
Db 417 SYPEFWE 423

RESULT 11
AROA_THEME STANDARD; PRT; 410 AA.
AC Q9WYI0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA OR TM0345.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF001715; AAD35431.1; -.
DR TIGR; TW0345; -.

```

[illegible]

Db	299	PMLOPITIEGSELVPKAIDELPVIALLCTQAVGTSTIKDAEBELKVKNRIRDTTADMNLNLL	358
Qy	369	GVCDEGETSVVVRGPDKGKLGNSAGAAVATHLDHRIAMSFVLGVLSNPVTVDATM	428
Db	359	GFELQPTNDGLIIH--PSEFKTNATDILT---DHRIGMMLAVACVLSSEPVKIQFDA	411
Qy	429	IATSFPEPM	437
Db	412	VNVSPGFL	420
RESULT	14		
AROA_HELPFJ			
ID	AROA_HELPFJ	STANDARD;	PRT; 429 AA.
AC	Q9ZKF7;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).		
GN	AROAR OR JHP0980.		
OS	Helicobacter pylori J99 (Campylobacter pylori J99).		
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;		
OC	Helicobacter.		
OX	NCBI_TaxID=85963;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99120557; PubMed=9923682;		
RA	Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,		
RA	Smith D.R., Noonan B., Guld B.C., deJonge B.L., Carmel G.,		
RA	Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,		
RA	Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,		
RA	Trust T.J.;		
RT	"Genomic sequence comparison of two unrelated isolates of the human		
RT	gastric pathogen Helicobacter pylori.";		
RL	Nature 397:176-180(1999)		
CC	-1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =		
CC	phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.		
CC	-1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN		
CC	THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).		
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).		
CC	-1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .		
CC	-----		
DR	EMBL; AE001527; AAD06557.1;		
DR	InterPro; IPR001986; EPSP_synthase.		
DR	Pfam; PF002075; EPSP_synthase; 1.		
DR	ProDom; PD001867; EPSP_synthase; 1.		
DR	PROSITE; PS00104; EPSP_SYNTHASE.1; 1.		
DR	PROSITE; PS00885; EPSP_SYNTHASE.2; 1.		
KW	Aromatic amino acid biosynthesis; Transferase; Complete proteome.		
SQ	SEQUENCE 429 AA; 47167 MW; 92724C4A2572741 CRC64;		
Query Match	22.8%;	Score 521.5;	DB 1; Length 429;
Best Local Similarity	33.68;	Pred. No. 1.5e-25;	
Matches	145; Conservative	73; Mismatches 166;	Indels 47; Gaps 13;
Qy	27	DKLSHSRFSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG-----DTWI	78
Db	10	DKLSHRAVIFSLAQKPCFVRNFMGECDSLSLEIAQLGAKVENTAKNSFKITPPTI	69
Qy	79	IDGVNGGLLAPEAPLDPEGNATCGRTMGVLGVYDFDSTFIGDASLTKRPWGRVLPNLR	138
Db	70	KE-----PNKILNCNSGTSMRLYSGLLSAOKGLFVLGSGNSLNARPKRIEPLK	120

DR TIGR: HP0401: -  
DR InterPro: IPR001986; EPSP\_synthase.  
DR Pfam: PF00275; EPSP\_synthase; 1.  
DR ProDom: PD001867; EPSP\_synthase; 1.  
DR PROSITE: PS00104; EPSP\_SYNTHASE\_1; 1.  
DR PROSITE: PS00885; EPSP\_SYNTHASE\_2; 1.  
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
SQ SEQUENCE 429 AA; 47240 MW; 19545753E081FDAE CRC64;

Query Match 22.7%; Score 518.5; DB 1; Length 429;  
Best Local Similarity 33.5%; Pred. No. 2.2e-25;  
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

QY 27 DKSTSHSFEGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKGG-----DTWI 78  
Db 10 DKSLSHRAVIFSLAQKPCFVRNFMAGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTI 69  
QY 79 IDGVNGGLLAPEAPLDFGNAATGCRMTGLVGVYDFDSTFIFGDASLTKRPMGRVNLPLR 138  
Db 70 KE-----PNKILNCNSGTTMRLYSGLLSAQKGLFVLSGDNLSNARPMKRIIEPLK 120  
QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPPTIT---YRVPNASAQVKSAVLLAGLNTPGITTVI 194  
Db 121 AFGAKILGREDNHFAPIVLG-----SPLKACHYESPIASAQVKSFAFSLSAQAQASTYK 176  
QY 195 EPIMTDRHTEKMLQGFANLTVETDADGVRTIR-LEGRGKLTGQVIDVPDGPSPSTAPPLV 253  
Db 177 ESELSRNHTEIMLSLADI---HNQDGVLKISPLE--KPLEAFDFTIANDPSSAFFAL 231  
QY 254 AALVPGSDVTILNLMNPTTGLILTLQEMGADIE--VINPRLAGGEDVADLRVRSSTL 311  
Db 232 ACAITPKSRLLKNVLLNPTRIEAFVLEKMGASIEYAIQSKDL---EMIGDITVEHAPL 288  
QY 312 KGVTVPEDRAPSMIDEXPIAVAAFAAGATVMNGLELRVKESDRLSAVANGKLNGVD 371  
Db 289 KAINIDQVIA-SLIDEIPALSIAMLFAGKSKMKNKDLRAKESDRKAVVSNFEKALGIE 347  
QY 372 CDEGETSLVVRGPDGKGL----GNASGAAVATHLDHRIAMSFVLMGLVSENPVTVDAT 427  
Db 348 CEEFEDGYVEGLEDISPLKQRFSEIKPPLIKSFNDHRIAMSFVLTIAL--PLEIDNLE 405  
QY 428 MIATSPFEEDIM 440  
Db 406 CANISFPQPKHLL 418

Search completed: August 15, 2002, 14:06:06  
Job time: 405 sec

QY 139 EMGVQVKS-EDGDRLPVTLRGPKTPPTITRYVPMASAQVKSAVLLAGLNTPGITTVIEPI 197  
Db 121 AFGAKILGREDNHFAPIVLGCPKA-CDYESPIASAQVKSFAFSLSAQAQGISAYKESE 179  
QY 198 MTRDHTKMLQGFANLTVETDADGVRTIR-LEGRGKLTGQVIDVPDGPSPSTAPPLVAAL 256  
Db 180 LSRNHTIMLSLGANI---QNQDGVLKISPLE--KPLESFDFETIANDPSSAFFALACA 234  
QY 257 LVPGSDVTILNLMNPTTGLILTLQEMGADIE--VINPRLAGGEDVADLRVRSSTLKGVT 315  
Db 235 ITPKSRLLKNVLLNPTRIEAFVLEKMGASIEYVIOSK--DLEVIGDIYIEHAPLKAIS 292  
QY 316 VPEDRAPSMIDEXPIAVAAFAAGATVMNGLELRVKESDRLSAVANGKLNGVDCEG 375  
Db 293 IDQVIA-SLIDEIPALSIAMLFAGKSKMKNKDLRAKESDRKAVVSNFEKALGIECEE 351  
QY 376 ETSLVVRGPDGKGLGNAS-----GAAVATHLDHRIAMSFVLMGLVSENPVTVD 425  
Db 352 EDGFIY-----EGUGDASQLKQHFSTKPKPIIKSFNDHRIAMSFVLTIAL--PLEIDN 403  
426 ATMIATSPPEF 436  
404 LECANISPTTF 414

RESULT 15  
ID AROA\_HELPY STANDARD; PRT; 429 AA.  
AC P56197;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
GN AROA OR HP0401.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter;  
OX NCBI\_Taxid=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey J.M.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
pylori.";  
RL Nature 388:539-547(1997).  
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate -  
phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE000556; AAD07470.1; -



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:51 ; Search time 51.67 Seconds  
(without alignments)  
1523.373 Million cell updates/sec

Title: US-09-464-099A-3  
Perfect score: 2288  
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP\_TREMBL\_19:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2104	92.0	455	16 Q92SV5	Q92sv5 rhizobium m
2	1861.5	81.4	480	2 Q9AGV2	Q9agv2 brucella ab
3	1789	78.2	452	16 Q98CC1	Q98cc1 rhizobium l
4	1133	49.5	443	16 Q9A2H2	Q9a2h2 pseudobacter
5	989.5	43.2	746	16 Q9H269	Q9hz69 pseudomonas
6	912.5	39.9	454	16 Q9PB21	Q9pb21 xylella fas
7	884	38.6	431	16 Q9KCA6	Q9kca6 bacillus ha
8	870	35.8	430	16 Q99ZB3	Q99zb3 streptococc
9	815.5	35.6	428	16 Q9A2A5	Q9a2a5 listeria in
10	781.5	34.2	428	2 Q9ANY6	Q9any6 enterococcu
11	610	26.7	432	16 Q99U25	Q99u25 staphylococ
12	470.5	20.6	207	2 Q9RHZ8	Q9rhz8 pseudomonas
13	470.5	20.6	439	17 Q9HQCL	Q9hqcl halobacteri
14	416.5	18.2	440	2 Q9I213	Q9i213 streptomyce
15	401.5	17.5	426	16 Q9KR80	Q9kr80 vibrio chol
16	378.5	16.5	427	2 Q93ED4	Q93ed4 yersinia ru

17	351.5	15.4	428	16 Q97KM2	Q97km2 clostridium
18	332.5	14.5	438	2 Q9K4A7	Q9k4a7 streptomyce
19	324.5	14.2	447	16 Q9K9D5	Q9k9d5 bacillus ha
20	306	13.4	516	10 Q946V0	Q946v0 dicliptera
21	299.5	13.1	511	10 Q93VK6	Q93vk6 oryza sativ
22	299	13.1	444	10 Q24566	Q24566 zea mays (m
23	298	13.0	516	10 Q946U9	Q946u9 dicliptera
24	296	12.9	521	10 Q9FVP6	Q9fvp6 arabidopsis
25	293.5	12.8	445	8 Q95AK0	Q95ak0 eleusine in
26	292.5	12.8	445	8 Q95AK1	Q95ak1 eleusine in
27	285	12.5	433	16 Q9JTT3	Q9jtt3 neisseria m
28	280	12.2	414	17 Q980I5	Q980i5 sulfolobus
29	275	12.0	433	16 Q9JYU1	Q9jyu1 neisseria m
30	261	11.4	408	17 Q96Y91	Q96y91 sulfolobus
31	260.5	11.4	427	17 Q9YC47	Q9yc47 aeropyrum p
32	229	10.0	332	2 Q9ZEQ0	Q9zeq0 actinobacil
33	223	9.7	410	17 Q978S3	Q978s3 thermoplasm
34	216.5	9.5	391	10 Q80428	Q80428 oryza sativ
35	204.5	8.9	448	2 Q9L1U5	Q9l1u5 streptomyce
36	201.5	8.8	446	2 Q9SON1	Q9son1 streptomyce
37	185.5	8.1	419	16 Q99Z78	Q99z78 streptococc
38	184.5	8.1	417	16 Q9K1Q9	Q9k1q9 neisseria m
39	184	8.0	418	16 Q97DD9	Q97dd9 clostridium
40	183.5	8.0	417	16 Q9JWS7	Q9jws7 neisseria m
41	182	8.0	428	16 Q9K6E5	Q9k6e5 bacillus ha
42	181	7.9	419	2 Q9SSW5	Q9ssw5 escherichia
43	179	7.8	347	10 Q9AT37	Q9at37 lolium rigi
44	179	7.8	423	16 Q927U1	Q927u1 listeria in
45	178.5	7.8	417	2 Q9EXE3	Q9exe3 mycobacteri

#### ALIGNMENTS

RESULT 1

Q92SV5 ID Q92SV5 PRELIMINARY; PRT; 455 AA.  
AC Q92SV5;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN (EC 2.5.1.19).  
OS Rhizobium melloti (Sinorhizobium melloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
ON NCBI\_TaxID=382;  
RX [1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21368234; PubMed=11474104;  
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubier F., Barnett M.J., Becker A., Boistard P., Bothe G., Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Guzy J., Gujral M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M., Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
RT "The composite genome of the legume symbiont Sinorhizobium melloti.";  
RL Science 293:668-672(2001).  
DR EMBL; AL591783; CAC41690.1; -;  
KW Transferase; Complete proteome.  
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 92.0%; Score 2104; DB 16; Length 455;  
Best Local Similarity 90.8%; Pred. No. 1.2e-119;  
Matches 413; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVPKDKISHRFEGGLASGETRITGLLEGEDVINTG 60

```
Db 1 MSHGASRRPATARKSSDGLKURIPGDKSISHRSMFGLAAGETRIIGLGEDVINTG 60
Qy 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Qy 121 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 180
Db 121 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 180
Qy 181 LLAGLNTPGITTVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGKGLTGQVID 240
Db 181 LLAGLNTPGITTVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGKGLTGQVID 240
Qy 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLTLTQEMGADIEVINPRLAGGED 300
Db 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLTLTQEMGADIEVINPRLAGGED 300
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP 420
Db 361 VADGLKNGVDCDEGASLVVRGPRDGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP 420
Qy 421 VTVDATMIATSPFPMGLMAGLAKIELSDTKAA 455
Db 421 VTVDATMIATSPFPMGLMAGLAKIELSDTKAA 455

RESULT 2
Q9AGV2 ID Q9AGV2 PRELIMINARY; PRT; 480 AA.
AC Q9AGV2;
AT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TremBLrel. 18, Last annotation update)
DE 5-ENOLPYRUVIL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).
GN AROA.
OS Brucella abortus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2308;
RA Gan T., Essenberg R.C.;
RT "Characterization of the araA gene of Brucella abortus and
RL Substitution of an araA mutant";
DR EMBL; AF326475; AAK27445.1; -
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 480 AA; 50634 MW; 3D55323944DA8C91 CRC64;

Query Match 81.4%; Score 1861.5; DB 2; Length 480;
Best Local Similarity 81.7%; Pred. No. 5,9e-105;
Matches 365; Conservative 35; Mismatches 44; Indels 3; Gaps 1;

Qy 1 MSHGASRRPATARKSSGLSGTGTIRIPGDKSISHRSMFGLAAGETRIIGLGEDVINTG 60
Db 31 MSHSACPKPATARRSQALTGEIRIPGDKSISHRSMFGLAAGETRIIGLGEDVINTG 90
Qy 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 91 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 150
```

```
Qy 121 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 180
Db 151 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 210
Qy 181 LLAGLNTPGITTVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGKGLTGQVID 240
Db 211 LLAGLNTPGITTVIEPIMTRDHTKMLQGFAGNLVETDADGVRTIRLEGKGLTGQVID 270
Qy 241 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLTLTQEMGADIEVINPRLAGGED 300
Db 271 VPGDPSSTAFPLVALLVPGSDVTILNLMNPTRTGLTLTQEMGADIEVINPRLAGGED 330
Qy 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 360
Db 331 VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 390
Qy 361 VANGKLNGVDCDEGETSLVVRGPRDGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP 420
Db 391 VARGLEANGVDCDEGETSLVVRGPRDGLGNASGAATAVATHLDRHRIAMSFVLMGLVSENP 447
Qy 421 VTVDATMIATSPFPMGLMAGLAKIELSDTKAA 447
Db 448 VTVDATMIATSPFPMGLMAGLAKIELSDTKAA 474

RESULT 3
Q98CC1 ID Q98CC1 PRELIMINARY; PRT; 452 AA.
AC Q98CC1;
AT 01-OCT-2001 (TremBLrel. 18, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TremBLrel. 18, Last annotation update)
DE 3-PHOSPHOSHIMIMATE 1-CARBOXYVINYLTRANSFERASE.
GN MLL5213.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51700.1; -
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR Transferase; Complete proteome.
KW SEQUENCE 452 AA; 47455 MW; 2B52983E3523B938 CRC64;

Query Match 78.2%; Score 1789; DB 16; Length 452;
Best Local Similarity 78.0%; Pred. No. 1.3e-100;
Matches 352; Conservative 33; Mismatches 62; Indels 4; Gaps 2;

Qy 1 MSHGASRRPATARKSSGLSGTGTIRIPGDKSISHRSMFGLAAGETRIIGLGEDVINTG 60
Db 1 MSHAAAKPATARKSQALSGTGTIRIPGDKSISHRSMFGLAAGETRIIGLGEDVINTG 60
Qy 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Db 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTGMLGVYDFDSTFI 120
Qy 121 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 179
Db 121 GDSASLTKRPMGRVNLPLRMGVQVKSDEGDRPLVTLRGPKTPPTTYRVPMAQAQVKSAY 179
```



Qy	4	GASSRPATARKSSGLSGTWRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAM	63
Db	5	GLKSAPGGA-----LRGIVRAPGDKSISHRSMTLIGALATGTTTVEGLDEGDDVLATARAM	59
Qy	64	QAMGARIRKEG-DTWIIDGVNGGLLAPAPLDIFGNAATGCRITMGLVGVYDFDSTFTGD	122
Db	60	QAFGARIRERGVGWRLE--KGCGPEPVDVDDCGNAGTGVRLIMGAAGFAMCATFTFGD	117
Qy	123	ASLTKRPMGRVNLPLRMGVQVKSDEGRDLPTVLTRGPKTPTTITYRVPMASAQVKSAVILL	182

[illegible]





Nordsiek G., Novella S., de Pablo B., Perez-Diaz J. C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Ramirez-Boland J.-A., Voss H., Weiland J., Cossart P.,  
RA "Comparative genomics of *Listeria* species.",  
RT Science 294:849-852(2001),  
RT EMBL; AL596170; CAC97267.1; -.  
DR L1stList; LIN02037; -.  
DR Complete proteome.  
SQ SEQUENCE 428 AA, 45994 MW, 157B48C091A68FEB CRC64;  
SQ

Query Match	35.6%	Score	815.5	DB	16	Length	428	
Best Local Similarity	40.8%	Pred.	No. 1e-41					
Matches	173	Conservative	83	Mismatches	155	Indels	13	Gaps
17	GLSGTVRIPGDKSISHRSPFMGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT	76						
	: : :       : :       : :     : :     : :     : :     : :     : :							
9	GLVGEITVPGDKSMHSRSMFGAIEAGKTVIRHFLRADDCLTGIRAKFALGVKIEETDEE	68						
	: : :       : :       : :     : :     : :     : :     : :     : :							
77	WIIDGVNGGGLLIAPEAPLDFGNATGCRUTMGLGVYDFDSTFFIGDASLTKRPMGRVLNP	136						
	: : :       : :       : :     : :     : :     : :     : :     : :							
69	IIVHGTSDDLKQAEGLPLDIGNSGTTRILMMGILAGRDFTVILGDSTAKRPMKRVMLP	128						
	: : :       : :       : :     : :     : :     : :     : :     : :							
137	LREMGVQVKSDDRL - PVTLRGEKPTPTITRVPMASQAQSVALLAGLNTPGTITVIE	195						
	: : :       : :       : :     : :     : :     : :     : :     : :							
129	LQEMGAKMKHKQGESEFAPISIGNQSLKMEYHMPVASQAQVSAIFAALQAEGETIIE	188						
	: : :       : :       : :     : :     : :     : :     : :     : :							
196	PIMTROHTEKMLGGGANITVETDAGVTIRLEGRKLTGQVIDVPGDPSATAPLVAA	255						
	: : :       : :       : :     : :     : :     : :     : :     : :							
189	KEKTRDTEHMIHQFGG - - - - ELEMDDL - TIRYKGGQKFQIGEMTPGVDSAAAFIVAG	243						
	: : :       : :       : :     : :     : :     : :     : :     : :							
256	LLVPGSDVTTLLVLMNPTRTGLTLTLOEMGADIEVLRNPLAGEDVADLRVRSSTLKGYT	315						
	: : :     : :     : :       : :       : :     : :     : :     : :							
244	LITPGSEIETHVGLNPTRTGIFDVVQMGGSLVVKDDSSRSTGKLAGTVVVKSSSELKGT	303						
	: : :     : :     : :       : :       : :     : :     : :     : :							
316	VPEDRAPSMTEPYIIIAVAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVDCDGG	375						
	: : :       : :       : :     : :     : :     : :     : :     : :							
304	IGGDIIPRLIDEIPVALLATQAEGTIIKDAELKVKETNRIDAVANELNKMGAIDITPT	363						
	: : :       : :       : :     : :     : :     : :     : :     : :							
376	ETSLVWGRPDGKGLGNASCAAAVATLHDHRIAMSFVLMG - LVSENPTVDDDATMIATSP	434						
	: : :       : :       : :     : :     : :     : :     : :     : :							
364	EDGLIIRGKTP - - - - - LHAANTSYGDHRIIGMLQIALLLVEDGDVELDRAFAVSVP	417						
	: : :       : :       : :     : :     : :     : :     : :     : :							
435	EFMD	438						
	: : :       : :       : :     : :     : :     : :     : :     : :							
418	TFE	421						
	: : :       : :       : :     : :     : :     : :     : :     : :							

RESULT	10	
Q9ANY6		
ID	Q9ANY6	PRELIMINARY; PRT; 428 AA.
AC	Q9ANY6;	
DT	01-JUN-2001	(TrEMBLrel. 17, Created)
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)
DT	01-OCT-2001	(TrEMBLrel. 18, Last annotation update)
DE	3-PROPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19).	
DN	ARO4	
OS	Enterococcus faecalis (Streptococcus faecalis).	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;	
OC	Enterococcus.	
OX	NCBI_TaxID=1351;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Og1RF;	
RA	Huycke M.M., Shepard L., Joyce W., Wise P., Moore D.R., Gilmore M.S.;	
RT	"Forme fruste respiration by Enterococcus faecalis produces	
RT	extracellular superoxide.";	
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF318277; AAG53678.1; -	
DR	InterPro; IPR001986; EFSP_syntase.	
DR	Pfam; PF00275; EFSP_syntase; 1.	
DR	ProDom; PD001867; EFSP_syntase; 1.	
DR	PROSITE; PS00104; EFSP_SYNTASE1; UNKNOWN_1.	

```

DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
SQ SEQUENCE 428 AA; 45715 MW; 9E0F4FEA4893CA95 CRC64;
  7;

Query Match 34.2%; Score 781.5; DB 2; Length 428;
Best Local Similarity 40.5%; Pred. No. 1.1e-39;
Matches 173; Conservative 79; Mismatches 152; Indels 23; Caps
7;

Oy 18 LSGVVRIPGDKSIHRSEMFEGGLASGETRIITGLLEGEDVINTGKAMGAMGARIRKEGDTW 77
Db 12 LQGLTLPVSDPDKSIHRSEMFEGGLASGETRIITGLLEGEDVINTGKAMGAMGARIRKEGDT 71
Oy 78 IIDGVNGGLLAPAPLDIFGNAATGCRITGLMGVGVYDFDSTFIDGSLTRKPMGRVLNPL 137
Db 72 TVEGRGFAGLKANTIDVNGSGTITRLMLGILAGCPFETRLAGDASIAKRPNNRVMPL 131
Oy 138 REMGVQVKS-EDGDRLPVTLRGKTPPTITRVPMASQAQVSAVLLAGLNTPGITTVIEP 196
Db 132 NQMAECOGVQQTFFPISIRGTQNLQPIDTVMVASQAQVSAVLLAGLNTPGITTVIEP 191
Oy 197 IMTRDHTKMLQGCANLTVETDADGVRTIRLEGKLTGQVIDVPGDPSPSTAFPLVAAL 256
Db 192 EKTRDHEEMIRQGGTLEV----DG-KKIMLTGQQLTGQVVVPGDISIAAFLVAGL 246
Oy 257 LVPGSDVILNVLMNPTRTGLTLQEMGADIEVINPLAGGEDVA-----DLVRVSTLK 312
Db 247 VVPDSEILLKNVGLNQTRTGILDVTKNMGSGVTILN-----EDEANHSGLLVKTSOLT 300
Oy 313 GVTVPDRAPSMIDEPYILAVAAFAEGATVMNGLEELURVKESDRLSAVANGLKLVDC 372
Db 301 ATEIGGAIPLRIDELPIALLATQATGTTIRDAELKVKETNRIDAVAKELTIGADI 360
Oy 373 DEGETSLVVRGPRDGKGLGNASAAVATHLDRHIAFSFVLMG-LVSENPVTVDDATMTAT 431
Db 361 PTFDGLLIHNGPT-----SLHGRVTSYGDHRIGMLQIAALLYKEGTVELDKAEAVSV 414
Oy 432 SPPEFMD 438
Db 415 SYPAFFD 421

RESULT 11
ID Q99U25 PRELIMINARY; PRT: 432 AA.
AC Q99U25;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 3-PROSPOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN AROA OR SAI297 OR SAV1464.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxId=158879, 158878;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES-S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL: AF003134; BAB4257.1; -.
DR EMBL: AF003362; BAB57626.1; -.
DR InterPro: IPR001986; EPSP_synthese

```

```
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; 1.  
DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.  
KW Transferrase; Complete proteome.  
SQ SEQUENCE 432 AA; 47068 MW; A42102057AD15C72 CRC64;  
  
Query Match 26.7%; Score 610; DB 16; Length 432;  
Best Local Similarity 34.0%; Pred. No. 2.8e-29;  
Matches 146; Conservative 80; Mismatches 181; Indels 22; Gaps 6;  
  
QY 15 SSGLSGTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKGG 74  
DB 10 SGPLAGEIEVPCDKSWTHRAIMLASLAEGTSNIYPLLEGEDCRRMTDIPRLGVGDIKDE 69  
  
QY 75 DTWIIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134  
DB 70 DKLVVNSPGYKAFKTPHOVLVTGNSGTTTRLLAGLLSGLGIESVLSDGVSIKGRPMGRVL 129  
  
QY 135 NPLREMGVQKSEGDRLPVTLRGPKTPPTTYRVPMSAQVKSAYVLLAGLNTPGITTVI 194  
DB 130 RPLKLMANIEGIDNYTPLIK-ESVTKGNYQMEVASAQVKSAYVLLAGLNTPGITTVI 188  
  
QY 195 EPIMTDRHTEKMQGF----GANTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249  
DB 189 ELDSVSNHTETFRFNPIEIERLSITTPDAIQHKPAD-----FHVPGDLSSAA 240  
  
QY 250 PFLVAALVPGSDVTILNVMNPTRLTLQEMGADIEVINPLAGGEDVADLRVR-S 308  
DB 241 FFIVAALITPESDVTHNVGINPTSGIIDIIVKMGNGNIOLFQ-OTTGAEPASIRIQYT 299  
  
QY 309 STLKGVTPEDRAPSWIDEPYILAAFAEGATVWNGLEELRVKESDRLSAVANGLKUN 368  
DB 300 PMLQPTITTEGELVPAIDELPVIALCTQAVGTSTIKDAEELKVKETNRIDTADMNLL 359  
  
QY 369 GVDCDEGETSLVVRGPRGDKGLGNASGAATHLDRHRTAMSLVGMVSENPVTVDATM 428  
DB 360 GFELQPTNDGLIIHSE-----FTNATVDSLTDRHIGMMLAVASLLSSEPVKIQFPA 413  
  
QY 429 IATSPPEFM 437  
DB 414 VNVSPFGL 422  
  
RESULT 12  
Q9RH28 PRELIMINARY; PRT; 207 AA.  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)  
DE 5-ENOLPYRUVYLISHIKMATE 3-P SYNTHASE (FRAGMENT).  
GN AROF.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-JM300;  
RX MEDLINE=92013931; PubMed=1919506;  
RA Fischer R.S., Zhao G., Jensen R.A.;  
RT "Cloning, sequencing, and expression of the P-protein gene (phea) of  
RT Pseudomonas stutzeri in Escherichia coli: implications for  
RT evolutionary relationships in phenylalanine biosynthesis.";  
RL J. Gen. Microbiol. 137:1293-1301(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JM300;  
RX MEDLINE=99298294; PubMed=10368439;  
RA Xie G., Bonner C.A., Jensen R.A.;  
RT "A probable mixed-function supraoperon in Pseudomonas exhibits gene  
organization features of both intergenomic conservation and gene  
shuffling.";  
RL J. Mol. Evol. 49:108-121(1999).  
DR EMBL: AF038578; AAD47363.1; -.  
DR InterPro: IPR001986; EPSP_synthase.  
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.  
FT NON_TER 207 207  
SQ SEQUENCE 207 AA; 21780 MW; 7231191C72A21D6B CRC64;  
  
Query Match 20.6%; Score 470.5; DB 2; Length 207;  
Best Local Similarity 51.5%; Pred. No. 2.8e-21;  
Matches 103; Conservative 22; Mismatches 70; Indels 5; Gaps 2;  
  
QY 12 ARKSSGLSTVRIPGDKSISHRSEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71  
DB 10 ANPGSGLTQGLRVPGDKSISHRSIMLGSIAEGTTEGFELEGEDALATXAQFRMGVVI- 68  
  
QY 72 KEG---DTWIIDGVNGGLLAPEAPLDFGNAATGRLTMGLVGVYDFDSTFIGDASLTKR 128  
DB 69 -EGPHQGRVTVHGVGLHGLQAPPPIYLGNSGTSMRLLAGLLAAQPFDTTLSGDASLTKR 127  
  
QY 129 PMGRVLNPLREMGVQKSEGDRLPVTLRGPKTPPTTYRVPMSAQVKSAYVLLAGLNT 188  
DB 128 PMNRVAKPLREMGAVTETAAGRPLTIRGGKKLSGMHYDMPMASAQVKSCLLAGLYAA 187  
  
QY 189 GITTVEIETMRDHTKMLQ 208  
DB 188 GKTSVTEPAPTRDHTERMQQ 207  
  
RESULT 13  
Q9HQC1 PRELIMINARY; PRT; 439 AA.  
AC Q9HQC1:  
DT 01-MAR-2001 (TremBLrel. 16, Created)  
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.  
GN PSC OR VNG1232G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium.  
OX NCBI_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005049; AAG19594.1; -.  
DR InterPro: IPR001986; EPSP_synthase.  
DR Pfam: PF00275; EPSP_synthase; 1.  
DR ProDom: PD001867; EPSP_synthase; 1.  
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.  
DR PROSITE: PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.  
KW Transferrase; Complete proteome.  
SQ SEQUENCE 439 AA; 44333 MW; 48CAD75A1F0CEB89 CRC64;  
  
Query Match 20.6%; Score 470.5; DB 17; Length 439;  
Best Local Similarity 31.3%; Pred. No. 7.8e-21;  
Matches 139; Conservative 76; Mismatches 190; Indels 39; Gaps 13;
```

Qy 16 SGLSTGVTRIPGDKSISHRSFMFGGLASGETRITIGLLEGEDVINTGKAMQAMGARIRKEGD 75  
Db 17 SRVRGRARAPSPKSYTHALLAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG 76  
Qy 76 TWIDGVNGGGLLAPAPLDFGNATGRLTMGLVGVYDFDSTIGDASLRKPMGRVLN 135  
Db 77 DWVVTGFSRPAI-PDAVIDCANSQTTMRLVTAALADAGTTLVLTGDSLRARPHGPLLD 135  
Qy 136 PLREMGVQVKSDEGD-RLPVTLRGPKTPTITYRVP-MASQVKSAVLLAGLNT-----P 188  
Db 136 ALSGLGGTARTRNGQAPLVVDGVSQGSVA--LPGVVSQFVALLMAGAVTGTET 193  
Qy 189 GTTIVIEPIMTRDHEKMLQFGANLTVETDADGVRTIRLEGRG-----KLTCQVIDVPGD 244  
Db 194 DLTTTELKSAFYVDITLDVDAFGVCA-ET-AAGYRV-----RGQAYAPSAEYAVPGD 246  
Qy 245 PSSATFPLVALLVPGSDVTLNLVLMNTRTG---LIITLQEMGADIETVPRLAGGEDV 301  
Db 247 FSSASYLLAAGALAAADGAAYVVEGMHPSAQGDAIIVDLVLERMGADID-----WDTES 299  
Qy 302 ADLRVRSSTLKGVTVPEDRAPSMIDEXPIAVAAAFAGATVMNGLEELRVKESDRLSAV 361  
Db 300 GVITVQSELSVEGVADTDPOLL---PTIAVLGNAADGTTITDAEHVRYKETDRVAA 356  
Qy 362 ANGLKLVGDCDEGETSLVVRGRDPKGLGNASGAATHLDHRTAMSLVNGLVSENPV 421  
Db 357 AESLSKLGASVEERDELVRG-----GDTLSGASVDGRGDRHRLVMAVAGLVADGET 411  
Qy 422 TVDDATMTATSPFPEMDLMAGLGA 445  
Db 412 TIAGSEHVDVSPFPDFEVLGLGA 435

RESULT 14  
Q9L213  
ID Q9L213 PRELIMINARY; PRT; 440 AA.  
AC Q9L213;  
DT 01-OCT-2000 (TremBrel. 15, Created)  
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)  
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.  
GN ARON.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Brown S.P., Harris D.;  
IL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmid and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL138598; CAB71266.1;  
DR InterPro; IPR001986; EFSP\_synthase.  
DR InterPro; IPR000408; RCC1.  
DR Pfam; PF00275; EFSP\_synthase.1.  
DR PRODOM; PD001867; EFSP\_synthase.1.  
DR PROSITE; PS00885; EFSP\_SYNTHASE.2; 1.  
DR PROSITE; PS00626; RCC1\_2; UNKNOWN\_1.  
KW transferase.

SQ SEQUENCE 440 AA; 46529 MW; 77B720F81398EB0D CRC64;  
Query Match 18.2%; Score 416.5; DB 2; Length 440;  
Best Local Similarity 28.0%; Pred. No. 1.5e-17;  
Matches 128; Conservative 78; Mismatches 198; Indels 53; Gaps 13;  
Qy 11 TARKSSGLSGTVTRIPGDKSISHRSFMFGGLASGETRITIGLLEGEDVINTGKAMQAMG--- 67  
Db 15 TIRLTGDEIVRVLSKSYNRYLAIASLSQETVIDNALLSDDTVVFSAIETFGHVT 74  
Qy 68 -----ARIKKEGTWIDGVNGGGLLAPAPLDFGNATGRLTMGLVGVYDFDSTFI 120  
Db 75 CDIDHATARIR-----VTPTGR-PMRAPSEDIVGGAGTPLRFLISMAGHADGTTIT 126  
Qy 121 GDASITTKPRMGRVLNPLREMGVQVKSDEGD-RLPVTLRGPKTPTITYRVPMSAQVKS 179  
Db 127 GNARMPGMDLLKALPALGDATAVRNGSPVVRVGGSGKATSGAVSQFTSS 186  
Qy 180 VLLAGLNPFGIT--TVIEPIMTRDHEKMLQFGANLTVETDADGVRTIRLEGRKLTGQ 237  
Db 187 LIINALRAQTDTEITISDDLVSQPVVEMTLAGL-AEMGVSVDRDGYRRFTVPSGOQARG 245  
Qy 238 VIDPQDSSSTAFPLVAA-----LLVPGSDVTLNLVLMNTRTGILTLQEMGADIEVI 291  
Db 246 QVTPEPDASGMSYFLAAAILQSRVVIQ-----IGASHQGDVHLVQALERMGRTVEV- 299  
Qy 292 NPLRAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEXPIAVAAAFAGATVMNGLEELR 351  
Db 300 -----GDD--SITVGGPLRGIDIDMEAMPDV--FSLAIVAAVAEGTTRITNIASLR 348  
Qy 352 VKESRLSAVANGKLVGDCDEGETSLVVRGRDPKGLGNASGAATHLDHRTAMSL 411  
Db 349 VKECRIAATVTELRKMGIDVEEHDAMYITG-----GTPHGAVIDTYDDHRIAMTFA 401  
Qy 412 VNLGVSENPVTVDDATMTATSPFPEMDLMAGLGA 448  
Db 402 IGLRTEG-VVIKDPCGVAKSPFAFWQTLDTLHPDLE 437  
RESULT 15  
Q9KR80  
ID Q9KR80 PRELIMINARY; PRT; 426 AA.  
AC Q9KR80;  
DT 01-OCT-2000 (TremBrel. 15, Created)  
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)  
DT 01-DEC-2001 (TremBrel. 19, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-  
DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)  
DE (EPSPS).  
GN VCI732.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae";  
RL Nature 406:477-483(2000).  
CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -  
CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.  
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE  
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

```
DR EMBL; AE004251; AAF94882.1; -.
DR TIGR; VC1732; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match      17.5%; Score 401.5; DB 16; Length 426;
Best Local Similarity 28.0%; Pred. No. 1.le-16;
Matches 125; Conservative 87; Mismatches 180; Indels 55; Gaps 15;

QY 18  LSGTVIPGDKSISHSRSMFGGLASGETRITGLLEGEDVINTKAMQAMGARIRKEGD-- 75
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 12  ISGEVNLPGSKSVSNRALLLAALASGTRTLNLLSDDIRHMLNALTCLGVNRYLSADKT 71

76  TWIIDGVNGGGLLAPAPLDGNAATGCRITMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
    |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 72  TCEVEGLGQAFHTTQPLEFLGNAGTAMRPLAAALCLGQGYVLTGEPMKERPIGHLVD 131

QY 136  PLREMGVQVKSEGDRLP-----VTLRGPKP---TPITYRVPMASAOVKS 179
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 132  ALRQAQAEIYLEQENFPLRIQCTGLQAGTVTIDGSISSQFLTAPLMSAPLAQGVTK 191

QY 180  VLLAGLNTFCITTVIEPIINTRDHTKMLQGFANLAVETDADGVRTIRLEGKGLT-GQV 238
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 192  IVGELVSKPVI-----DITLHIMEQFGVQV--INHDIQEFVIPAGQSVVSPGQF 238

QY 239  IDVPGPSSATPLVAALLVPGSDVTILNLMNPTRTGLIL--TLQEMGADIEVINPRLA 296
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 239  L-VEGDASSASY-FLAAAIKGEVKTGIGKNSIQGDIQFADALEKMGQAIE----- 289

QY 297  GGEDVADLRVRSSTLKGTVVPEDRAPSMIDEXYPILAVAAAFAGATVMNGLEELRVKESD 356
    |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 290  WGDDY--VIARGELNAVLDLDFNHIP---DAAMTIATTALFAKGTTAIRNVYWRVKETD 344

QY 357  RLSSAVANGLKNGVDCDEGETSLVYVGRPDGKGLGNASGAAVATHLDHRIAMSFYMGVLV 416
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 345  RLAAAMATELRKVGATVEEGEDFIVI--TPPTKLI----HAAIDYDDHRMAMCFSLVAL- 397

QY 417  SENPVTVDATMTATSPFEFMDLMAGL 443
    |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 398  SDPPVTINDPKCTSKTFPDYDKFAQL 424
```

Search completed: August 15, 2002, 14:05:43  
Job time: 412 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:41 ; Search time 56.25 Seconds  
 (without alignments)  
 898.464 Million cell updates/sec

Title: US-09-464-099A-3  
 Perfect score: 2288  
 Sequence: 1 MSHCASSRPATARKSSGLSG.....FMDLMAGLAKIELSDTKAA 455

Scoring table:  
 BLOSUM62  
 Gapop 10.0 , Gapext 0.5

searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

1:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	455	13 AAR22300	Class II EPSPS enz
2	2288	100.0	455	18 AAW34683	Class II EPSP synt
3	2288	100.0	455	18 AAW24474	Class II EPSPS for
4	2288	100.0	455	19 AAW71609	Agrobacterium sp.
5	2288	100.0	455	22 AAE05053	Agrobacterium sp.
6	2282	99.7	455	13 AAR26449	CP4-EPSPS. Synthe
7	2282	99.7	455	19 AAW39426	CP4-EPSPS protein.
8	2282	99.7	527	22 AAW52214	EPSPS SEQ ID NO 3.
9	1906.5	83.3	449	13 AAR22301	Class II EPSPS enz
10	1906.5	83.3	449	13 AAR22302	Class II EPSPS enz
11	1906.5	83.3	449	18 AAW34684	Class II EPSP synt

12	1906.5	83.3	449	18 AAW34685	Class II EPSP synt
13	1906.5	83.3	449	18 AAW24479	Class II EPSPS use
14	1906.5	83.3	449	18 AAW24480	Class II EPSPS use
15	1906.5	83.3	449	19 AAW71611	Pseudomonas sp. st
16	1906.5	83.3	449	19 AAW71610	Achromobacter sp.
17	1906.5	83.3	449	22 AAE05054	Achromobacter sp.
18	1906.5	83.3	449	22 AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18 AAW34688	Class II EPSP synt
20	867.5	37.9	447	18 AAW24487	Class II EPSPS for
21	867.5	37.9	447	19 AAW71619	Synechocystis sp.
22	867.5	37.9	447	22 AAE05072	Synechocystis sp.
23	806	35.2	443	18 AAW34689	Class II EPSP synt
24	806	35.2	443	18 AAW24488	Class II EPSPS for
25	806	35.2	443	19 AAW71620	Dichelobacter nodo
26	806	35.2	443	22 AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22 AAB48177	S. pneumoniae aroA
28	803.5	35.1	427	22 AAB37094	Streptococcus pneu
29	803.5	35.1	427	22 AAB48500	Streptococcus pneu
30	798.5	34.9	427	20 AAW97388	5-enolpyruvylshikl
31	798.5	34.9	427	21 AAY56504	Streptococcus pneu
32	798.5	34.9	427	22 AAB31158	A 5-enolpyruvylsh
33	785	34.3	428	18 AAW34686	Class II EPSP synt
34	785	34.3	428	18 AAW24481	Class II EPSPS use
35	785	34.3	428	19 AAW71617	Bacillus subtilis
36	785	34.3	428	22 AAE05070	Bacillus subtilis
37	704.5	30.8	415	20 AAW97389	Protein encoded by
38	704.5	30.8	415	21 AAY56505	Streptococcus pneu
39	704.5	30.8	415	22 AAB31159	A 5-enolpyruvylsh
40	704.5	30.8	415	22 AAB48178	S. pneumoniae aroA
41	704.5	30.8	415	22 AAB37095	Streptococcus pneu
42	704.5	30.8	415	22 AAB48501	Protein encoded by
43	577	25.2	430	18 AAW34687	Class II EPSP synt
44	577	25.2	430	18 AAW24482	Class II EPSPS use
45	577	25.2	430	19 AAW71618	Staphylococcus aur

## ALIGNMENTS

RESULT 1

AAR22300

ID AAR22300 standard; Protein; 455 AA.

XX

AC AAR22300;

XX

DT 03-AUG-1992 (first entry)

XX

Class II EPSPS enzyme.

XX

DE

XX

KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;

KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

XX

OS Agrobacterium sp. strain CP4.

XX

PN WO9204449-A.

XX

PD 19-MAR-1992.

XX

PF 28-AUG-1991; 91WO-US06148.

XX

PR 31-AUG-1990; 90US-0576537.

XX

(MONS ) MONSANTO CO.

PA Barry GF, Kishore GM, Padgett SR;

XX WPI; 1992-114356/14.

XX DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate

PT synthase - for producing plants and bacteria tolerant to

XX glyphosate herbicides

XX

PS Disclosure; Fig 3; 148pp; English.

XX	Key	Location/Qualifiers
FH	Region	200...204
FT		/label= characteristic_region
FT		/note= "see AAW34690"
FT		26..29
FT	Region	/label= characteristic_region
FT		/note= "see AAW34691"
FT		

[illegible][illegible]

QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPNTRGLILTLQEMGADIEVINPRLAGGED 300  
|||||  
Db 241 vpgdpsstafplvaallvpgsdvtilnvlmpntrglilqlqemgadievlnprlagged 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEXYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
|||||  
Db 301 vadlrsvrstlkgvtvpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360  
QY 361 VANGILKLVGDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420  
|||||  
Db 361 vangiklvngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420  
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA 455  
|||||  
Db 421 vtvdattmiatsfpefmdlmaglakielstdtkaa 455

## RESULT 3

AAW24474  
AAW24474 standard; Protein; 455 AA.

NC AAW24474;

DT 02-OCT-1997 (first entry)

XX Class II EPSPs for glyphosate resistant plant production.

XX 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; Class II;  
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;  
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;  
KW tobacco.

XX Agrobacterium sp. strain CP4.

OS AAW24474

PN US5633435-A.

XX 27-MAY-1997.

XX 31-AUG-1990; 90US-0576537.

XX 13-SEP-1994; 94US-0306063.

XX 31-AUG-1990; 90US-0576537.

XX 28-AUG-1991; 91US-0749611.

XX (MONS ) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

XX WPI; 1997-297418/27.

XX N-PSDB; AAT77313.

XX New isolated 5-enolpyruvylshikimate-3-phosphate synthase gene -

XX used for transforming plants to produce plants which are tolerant to

XX glyphosate herbicide

XX Claim 1; Column 57-60; 154pp; English.

PS AAW24474 shows the sequence of a class II 5-enolpyruvylshikimate-3-

CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant

CC to glyphosate herbicides. EPSPS and sequences encoding it are used for

CC the production of herbicide resistant (glyphosate-tolerant) plants

CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed

CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,

CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.

XX Sequence 455 AA;

Query Match

Best Local Similarity 100.08; Score 2288; DB 18; Length 455;

Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFPMFGGLASGETRITCLLEGEDVINTG 60

Db 1 mshgassrpataarkssglsgtvlripgdkshrsfpmfgglasgetritcllegedvintg 60  
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPDLDFGNAATGCRITMGLVGVYDFDSTFI 120  
|||||  
Db 61 kamqamgarirkegdtwiiidgvnggllapdlfdgnaatgcriltmglvgvydfdstfi 120  
QY 121 GDASTLTKRPMGRVNLPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYRVPMAQAQVKSAY 180  
|||||  
Db 121 gdasitlkrpmgrvnlplremgvqvkssedgdrilpvtlrgpktpitryrvmasaqvksay 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFGANLTVETDADGVRTIRLGRGLTQVID 240  
|||||  
Db 181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltqgvid 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNVLNMPNTRGLILTLQEMGADIEVINPRLAGGED 300  
|||||  
Db 241 vpgdpsstafplvaallvpgsdvtilnvlmpntrglilqlqemgadievlnprlagged 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEXYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
|||||  
Db 301 vadlrsvrstlkgvtvpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360  
QY 361 VANGILKLVGDCDEGETSLVVRGPDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420  
|||||  
Db 361 vangiklvngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420  
QY 421 VTVDATMIATSPFPEFMDLMAGLAKIELSDTKAA 455  
|||||  
Db 421 vtvdattmiatsfpefmdlmaglakielstdtkaa 455

## RESULT 4

AAW71609

ID AAW71609 standard; Protein; 455 AA.

XX AC AAW71609;

DT 14-DEC-1998 (first entry)

XX Agrobacterium sp. strain CP4 Class II EPSPs.

DE 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPs enzyme;  
KW glyphosate herbicide; transformed bacteria; class I EPSPs enzyme;  
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;  
KW N-phosphonomethylglycine.

XX Agrobacterium sp.

OS US5804425-A.

PN 08-SEP-1998.

XX 07-APR-1997; 97US-0833485.

XX 13-SEP-1994; 94US-0306063.

XX 31-AUG-1990; 90US-0576537.

XX 28-AUG-1991; 91US-0749611.

XX 07-APR-1997; 97US-0833485.

XX (MONS ) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

XX WPI; 1998-505657/43.

XX N-PSDB; AAW58009.

XX Glyphosate resistant 5-enolpyruvylshikimate-3-phosphate synthase -

XX useful for characterisation of the enzyme to determine inhibition

XX data values

XX Claim 3; Fig 3; 152pp; English.

CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the  
 CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),  
 CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,  
 CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,  
 CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,  
 CC produced by recombinant methods, can be used in kinetic studies to  
 CC determine Ki and Km values of the enzyme for its characterisation. The  
 CC enzyme is normally used for the production of 5-enolpyruvyl-3-  
 CC phosphoshikimate in plants, and most forms of the enzyme are  
 CC inhibited by N-phosphonomethylglycine (glyphosate) herbicides.  
 CC Inhibition data enables more accurate values of concentrations of  
 CC herbicide to be used when growing the plant without being detrimental to  
 CC it. This enables the plant to be grown in the presence of the herbicide,  
 CC being used to inhibit the growth of undesired plants. The present  
 CC sequence represents a Class II EPSPS from bacterial isolate  
 CC Agrobacterium sp. strain CP4.  
 XX  
 XX Sequence 455 AA;

Query Match 100.0%; Score 2288; DB 19; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-185;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVRIPODKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60  
 DB 1 mshgassrpataarkssglsgrtvpdkksishrsfmfgglasgetritgllegedvintg 60  
 QY 61 KAMQANGARIRKEDGTWIIDGVNGGLLAPLDFGNATGCRITMGLVGYVDFDSTFI 120  
 DB 61 kamqangarirkegdwtwiiidgvnggllapeapldfignatgcriltmglvgyvdfdstfi 120  
 QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPIRYVPMSAQVKSAY 180  
 DB 121 gdasltkrpgrvnlplremgvqkseddgdrilpvtlrgpktpitpityrvmasaqvksav 180  
 QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQVID 240  
 DB 181 llaglntpgittviepiimtrdhtekmlqgfngantvetdadvrtirlgrgkltqgvid 240  
 QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNLMNPTRTGLILTLQBMGADIEVINPRLAGGED 300  
 DB 241 vpgdpsstaplvaallvpgsdvtiilnlnmptrtgliltlqemgadvieinprlagged 300  
 QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPLAVAAFAEGATVMNGLEELRVKESDRLSA 360  
 DB 301 vadlrvrssltkgvtpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360  
 QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSLVMGLVSENP 420  
 DB 361 vangklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420  
 QY 421 VTVDATMTATSPFMDLMAGLAKIELSDTKAA 455  
 DB 421 vtvdattmtatspfmdlmaglakielstdtkaa 455

RESULT 5  
 AAE05053  
 ID AAE05053 standard; Protein; 455 AA.  
 XX  
 XX AAE05053;  
 AC  
 DT 10-SEP-2001 (first entry)  
 XX Agrobacterium sp. strain CP4 class II EPSPS protein.  
 XX 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;  
 XX transformed bacteria; transgenic plant; herbicide.  
 XX Agrobacterium sp. CP4.  
 OS  
 XX US6248876-B1.  
 PN

XX 19-JUN-2001.  
 XX 20-AUG-1998; 98US-0137440.  
 XX 13-SEP-1994; 94US-0306063.  
 XX 07-APR-1997; 97US-0833485.  
 XX 31-AUG-1990; 90US-0576537.  
 XX 28-AUG-1991; 91US-0749611.  
 XX (MONS ) MONSANTO CO.  
 XX Barry GF, Kishore GM, Padgett SR, Stallings WC;  
 XX WPI; 2001-407326/43.  
 XX N-PSDB; AAD09754.  
 XX  
 XX DNA probe capable of use in a polymerase chain reaction for identifying  
 XX the presence of a target genomic DNA encoding a  
 XX 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme -  
 XX Claim 3; Fig 3; 152pp; English.  
 XX  
 XX The present invention relates to a DNA probe capable of use in a  
 XX polymerase chain reaction for identifying the presence of a target  
 XX genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)  
 XX enzyme referred to as class II EPSPS which is tolerant to  
 XX glyphosate. EPSPS genes are useful in producing transgenic bacteria  
 XX and transgenic plants which are tolerant to glyphosate herbicide. The  
 XX probe is useful for identifying the presence of a target genomic DNA  
 XX encoding a EPSPS enzyme. The present sequence is Agrobacterium sp.  
 XX strain CP4 class II EPSPS protein.  
 XX Note: The present sequence, SEQ ID NO: 3 is stated as amino acid  
 XX sequence throughout the specification. However, it is referred as probe  
 XX in claim 3 of the specification.  
 XX  
 XX Sequence 455 AA;

Query Match 100.0%; Score 2288; DB 22; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-185;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSSGLSGTVRIPODKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60  
 DB 1 mshgassrpataarkssglsgrtvpdkksishrsfmfgglasgetritgllegedvintg 60  
 QY 61 KAMQANGARIRKEDGTWIIDGVNGGLLAPLDFGNATGCRITMGLVGYVDFDSTFI 120  
 DB 61 kamqangarirkegdwtwiiidgvnggllapeapldfignatgcriltmglvgyvdfdstfi 120  
 QY 121 GDASLTKRPMGRVNLPLREMGVQKSEDDGRLPVTLRGPKTPTPIRYVPMSAQVKSAY 180  
 DB 121 gdasltkrpgrvnlplremgvqkseddgdrilpvtlrgpktpitpityrvmasaqvksav 180  
 QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRKLTQVID 240  
 DB 181 llaglntpgittviepiimtrdhtekmlqgfngantvetdadvrtirlgrgkltqgvid 240  
 QY 241 VPGDPSSTAPFLVAALLVPGSDVTILNLMNPTRTGLILTLQBMGADIEVINPRLAGGED 300  
 DB 241 vpgdpsstaplvaallvpgsdvtiilnlnmptrtgliltlqemgadvieinprlagged 300  
 QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPLAVAAFAEGATVMNGLEELRVKESDRLSA 360  
 DB 301 vadlrvrssltkgvtpedrapsmideypilavaafaegatvmngleelrvkesdrlsa 360  
 QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATVATHLDRHRIAMSLVMGLVSENP 420  
 DB 361 vangklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420  
 QY 421 VTVDATMTATSPFMDLMAGLAKIELSDTKAA 455  
 DB 421 vtvdattmtatspfmdlmaglakielstdtkaa 455

Db 421 vtvdattmatsfpefmdlmaglakielstkaa 455

RESULT 6

AAW39426 standard; Protein; 455 AA.

XX AAR26449  
AC AAR26449;  
XX 28-JAN-1993 (first entry)  
XX CP4-EPSPS.  
XX pMON11030; CTP2; CP4; EPSPS; chloroplast transit peptide;  
KW 5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;  
KW Pseudomonas chloroaphis 6G5.  
XX Synthetic.  
XX WO9212249-A.  
XX 23-JUL-1992.  
XX 17-DEC-1991; 91WO-0509437.  
XX 26-DEC-1990; 90US-0632440.  
XX (MONS ) MONSANTO CO.  
XX Kishore GM, Klee HJ;  
XX WPI; 1992-284334/34.  
XX N-PSDB; AAQ27201.

Delaying fruit ripening and senescence in plants - by controlling ethylene prodn., pref. by expression of 1-amino:cyclopropane-1-carboxylic acid deaminase  
XX  
XX Disclosure; Page 64-66; 110pp; English.  
XX The sequences given in AAR26448 and AAR26449 are encoded by genes which were used in the construction of an expression plasmid pMON11030 which was used to transform petunia plants. This plasmid contained the genes for chloroplast transit peptide (Ctp2) (AAR26448) and the CP4 synthetic 5-enolpyruvyl-3-shikimate phosphate synthetase (EPSPS) gene (AAR26449) which is capable of conferring resistance to glyphosate. The plasmid also contained the 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase gene from Pseudomonas chloroaphis (see also AAQ27199). The transformed plants had ethylene levels reduced to about one half that of the control, untransformed plants. It is expected that such plants will show reduced senescence of flowers and leaves when compared to untransformed plants.  
XX  
XX Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 13; Length 455;  
Best Local Similarity 99.8%; Pred. No. 2e-184;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSHGASSRPTARKSGLSGTVRTPGDKSISHRSFMFGGLASGETRTITGLLEGEDVINTG 60  
Db 1 mlhgassrptarksglsgrvripgdkshrsfmfgglasgetritgllegedvintg 60  
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPAPLDGFAATGCRLLTMGLVGVYDFDSTFI 120  
Db 61 kamqamgarirkegdtwiidgvnggllapeapldfnaatgcrlltmglvgvydfdstfi 120  
QY 121 GDASLTRPMGRVNLPREMGVQVKSDEGRLPVTILRGPKTPPTITRVPMASQVKSVA 180  
Db 121 gdasltrpmgrvnlpremgvqvksedgrlpvtlrgpktptityrvpmasqvksav 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRGLTGQVID 240

Db 181 lllaglnptgittviepimtrdhtekmlqggfanltvetdadvrtirlegrgkltgqvld 240  
QY 241 VPGDPSSTAPFLVAALLVPGSDVYTLNVLNMPRTGTLILTLQEMGADIEVINPRLAGGED 300  
Db 241 vpgdpsstapflvaallvpgsdvtilnvlmptrtgitlilqlqemgadievlnprlagged 300  
QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPVTLAVAAAFAGATVMNGLEELRVKESDRLSA 360  
Db 301 vadlrvsrstlkgvtvpedrapsmideypllavaaaafagatvmngleelrvkesdrisa 360  
QY 361 VANGKLNGVDCDEGETSLVVRGRPDGKGLGNASGAATVATHLDRHRIAMSFVLMGLVSEN 420  
Db 361 vangklngvdcdegetslvvrgrpdkgkignasgaatvathldhriamsvflmglvsenp 420  
QY 421 VTVDATMIATSPFPEFMDLMAGLGAKEIELSSTKAA 455  
Db 421 vtvdattmatsfpefmdlmaglakielstkaa 455

RESULT 7

AAW39426 standard; Protein; 455 AA.

XX AAW39426;  
XX 19-MAY-1998 (first entry)  
XX CP4-EPSPS protein.  
XX 1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;  
KW fruit ripening; ethylene production; glyphosate resistance;  
KW 5-enolpyruvyl-3-shikimate phosphate synthase; EPSPS; CP4.  
XX Synthetic.  
XX US5702933-A.  
XX 30-DEC-1997.  
XX 06-NOV-1995; 95US-0553943.  
XX 17-DEC-1991; 91US-0809457.  
XX 26-DEC-1990; 90US-0632440.  
XX 06-NOV-1995; 95US-0553943.  
XX (MONS ) MONSANTO CO.  
XX Kishore GM, Klee HJ;  
XX WPI; 1998-076419/07.  
XX N-PSDB; AAV09719.

Production of plants with delayed ripening - using DNA encoding 1-amino:cyclopropane-1-carboxylic acid deaminase  
XX  
XX Example 9; Fig 21; 56pp; English.  
XX This sequence represents the synthetic CP4 5-enolpyruvyl-3-shikimate phosphate synthase (EPSPS) gene which is capable of conferring resistance to glyphosate and is used in a novel method for producing fruit-bearing plants with delayed ripening. The method involves the expression of a 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase gene in a plant at a level sufficient to reduce ethylene production in the fruit.  
XX  
XX Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 19; Length 455;  
Best Local Similarity 99.8%; Pred. No. 2e-184;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSHGASSRPATARKSGSLSGTVIRIPGDKSISHRSFMPFGLASGETRITGCLLEGEDVINTG 60  
Db 1 mhgassrpataarksgslsgstvripdgksishrsfmgfglasgetritgcllegedvintg 60  
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATCRLTMGLVGYDYDFDSTFI 120  
Db 61 kamqamgarirkegdtwiidgvnggllapeapldfgnaatcrltmglvgydydfdstfi 120  
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTTYRVPMSAQVKSAY 180  
Db 121 gdasltkrpmgrvnlplremgvqvksegdrlpvtlrgpktpttyrvpmsaqvksav 180  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLGRGKLTGOVID 240  
Db 181 llaglntpgittviepiimtrdhtkmlqgfanltvetdadvrtirlegrgkltgqvvid 240  
QY 241 VFGDPSSTAFPLVAALLVPGSDVTILNVLNPTRTGLIILTQEMGADIEVINPRLAGGED 300  
Db 241 vfgdpsstaflvaallvpgsdvtlinvlnpstrtgliltqemgadievlnprlagged 300  
QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLLELRVKESDRLSA 360  
Db 301 vadlrvrstlkgvtvpedrapsmideypilavaafaegatvmnglleelrvkesdrlsa 360  
QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420  
Db 361 vangklngvdcdegetsllvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420  
QY 421 VTVDATMIATSPPEPMDLMAGLAKIELSDTKAA 455  
Db 421 vtvdattmiatspfePMDlmaglakielstdtkaa 455  
RESULT 8  
AA052214  
ID AA052214 standard; Protein; 527 AA.  
XX  
AC AA052214;  
DT 12-FEB-2002 (first entry)  
XX  
DE EPSPS SEQ ID NO 3.  
KW Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;  
KW 5'-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;  
KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;  
KW chloroplast transit peptide.  
OS Petunia hybrida.  
XX Agrobacterium sp.  
PN JP2001190168-A.  
PD 17-JUL-2001.  
XX  
XX 27-OCT-2000; 2000JP-0328811.  
XX  
XX 29-OCT-1999; 99JP-0310244.  
XX  
XX (SUMO ) SUMITOMO CHEM CO LTD.  
XX  
XX WPI; 2001-605307/69.  
DR N-PSDB; ABA02854.  
XX  
XX New herbicide-resistant plant -  
XX  
XX Examples; Page 46-48; 66pp; Japanese.  
XX  
XX The invention relates to a transgenic plant which shows resistance to a  
CC herbicide in an amount inhibiting natural  
CC 5'-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant  
CC and having at least one enzymatic activity selected from:  
CC (1) EPSPS activity different from natural EPSPS activity of the plant or

CC (2) glyphosate oxidoreductase (GOX) activity different from the natural  
CC GOX activity of the plant; and in which a gene encoding a protein having  
CC the following properties: (a) combines specifically with a substance  
CC participating to the herbicidal activity a herbicide of  
CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no  
CC denaturing activity on a substance to which said protein combines  
CC specifically; and (c) contains substantially no framework region of the  
CC variable region of immunoglobulin. The present sequence is that of the  
CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium  
CC sp. strain CP4 EPSPS gene.  
XX  
SQ Sequence 527 AA;  
Query Match 99.7%; Score 2282; DB 22; Length 527;  
Best Local Similarity 99.8%; Pred No. 2, 5e-184;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSHGASSRPATARKSGSLSGTVIRIPGDKSISHRSFMPFGLASGETRITGCLLEGEDVINTG 60  
Db 73 mhgassrpataarksgslsgstvripdgksishrsfmgfglasgetritgcllegedvintg 132  
QY 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATCRLTMGLVGYDYDFDSTFI 120  
Db 133 kamqamgarirkegdtwiidgvnggllapeapldfgnaatcrltmglvgydydfdstfi 192  
QY 121 GDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTTYRVPMSAQVKSAY 180  
Db 193 gdasltkrpmgrvnlplremgvqvksegdrlpvtlrgpktpttyrvpmsaqvksav 252  
QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFANLTVETDADGVRTIRLGRGKLTGOVID 240  
Db 253 llaglntpgittviepiimtrdhtkmlqgfanltvetdadvrtirlegrgkltgqvvid 312  
QY 241 VFGDPSSTAFPLVAALLVPGSDVTILNVLNPTRTGLIILTQEMGADIEVINPRLAGGED 300  
Db 313 vfgdpsstaflvaallvpgsdvtlinvlnpstrtgliltqemgadievlnprlagged 372  
QY 301 VADLRVRSSTLKGTVTPEDRAPSMIDEYPIILAVAAFAEGATVMNGLLELRVKESDRLSA 360  
Db 373 vadlrvrstlkgvtvpedrapsmideypilavaafaegatvmnglleelrvkesdrlsa 432  
QY 361 VANGKLINGVDCDEGETSLVVRGPRDGKGLGNASGAATAVATHLDHRIAMSFVLMGLVSENP 420  
Db 433 vangklngvdcdegetsllvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 492  
QY 421 VTVDATMIATSPPEPMDLMAGLAKIELSDTKAA 455  
Db 493 vtvdattmiatspfePMDlmaglakielstdtkaa 527  
RESULT 9  
AA022301  
ID AAR22301 standard; Protein; 449 AA.  
XX  
AC AAR22301;  
DT 03-AUG-1992 (first entry)  
XX  
DE Class II EPSPS enzyme.  
XX  
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;  
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.  
XX  
OS Achromobacter sp. strain LBAA.  
XX  
XX WO9204449-A.  
XX  
PD 19-MAR-1992.  
XX  
XX 28-AUG-1991; 91WO-US06148.  
XX  
XX 31-AUG-1990; 90US-0576537.  
PR

```

XX (MONS ) MONSANTO CO.
XX
XX Barry GF, Kishore GM, Padgett SR;
XX
XX WPI; 1992-114356/14.
XX
XX DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
XX synthase - for producing plants and bacteria tolerant to
XX glyphosate herbicides
XX
XX Disclosure; Fig 5; 148pp; English.
XX
XX The sequence is that of the Class II 5'-enolpyruvylshikimate-3
XX phosphate synthase enzyme (EPSPS) of Achromobacter sp. strain LBAA
XX It is used to create glyphosate resistant plants or seeds which
XX can be planted in a field of crops to selectively control weeds.
XX The crops selected for are e.g. corn, wheat, rice, oilseed rape,
XX tobacco and alfalfa. This provides a cost effective, environmentally
XX compatible weed control device. See also AAR22300 and AAR22302.
XX
XX Sequence 449 AA;
XX
XX Query Match 83.3%; Score 1906.5; DB 13; Length 449;
XX Best Local Similarity 83.1%; Pred. No. 1.1e-152;
XX Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps
XX
XX QY 1 MSHGASSRPATARKSSCLGCTVRIPGDKSISHRSFMPGGLASGETRITGLLEGEDVINTG 60
XX Db 1 mshsaspkpatarsealtgeiripgdkshrsfmgfllasgetritgilegedvintg 60
XX QY 61 KAMQAMGARIRKEGDTWIDGVNGGLLAPPEALDFGNAAATGRLTMGLGVYDFDSTFI 120
XX Db 61 ramqamgakirkegdwivngvgcllqpeaaldfgnagtgartlmglvgtymktsfi 120
XX QY 121 GDASLTRKPRGVRVNLPREMGVQVQKSEGDRLPVTLRGPKTPTITYRVPMSAQVKSAY 180
XX Db 121 gdaslskrpmgrvlnpremgvqveaadgrmpitligpktaupityrvpmasagvksav 180
XX QY 181 LLAGLNTPGTWTVEPIWTRDHTKMLQGFGANLTVETDADGVRTLRGGRKLTGQVID 240
XX Db 181 llaglnpcgvttvlepvmtrdhckmlqgfgadltvtdkgvrhritggqklvgqtd 240
XX QY 241 VPGDPSSTAPFLVAALVPGSDVTILNVLNPNRTGLILTQEMGADIEVNPRLAGGED 300
XX Db 241 vpgdpsstapflvaalivesgdvtrnvlmptrtgililqemgadielvarlagged 300
XX QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAAVAAFAEGATVYNGLEELRVKESDRLSA 360
XX Db 301 vadlrvaslklgvvpperapsmiddeypvllaasfaegtvmgdldelrvkesdrlla 360
XX QY 361 VANGKLINGVDCOGEETSLVVRGPRDCKGLGNASGAAVATHLDHRAMSLVLMGLVSENP 420
XX Db 361 vargleangvdcetegmsltvrgprdgkgig---ggvtvathldhriamsflvmglaekp 471
XX QY 421 VTVDATMIATSPFEFMDLMAGLCAKIELS 450
XX Db 418 vtvdsmiatsfepfndmmpgigakiels 447
XX
XX RESULT 10
XX AAR22302
XX ID AAR22302 standard; Protein; 449 AA.
XX AC AAR22302;
XX XX
XX DT DT
XX XX
XX DE DE
XX XX
XX KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants
XX KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

```

[illegible]

ID XX AAW34684 standard; Protein; 449 AA.

AC AAW34684;

XX 17-FEB-1998 (first entry)

DT Class II EPSP synthase (EPSPs) from *Achromobacter* sp. strain LBAA.

DE 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; glyphosate;

XX shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;

XX 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;

XX glyphosate resistance gene; glyphosate-tolerance; promoter.

OS *Achromobacter* sp. strain LBAA.

XX Key Location/Qualifiers

XX Region 200..204

FT /label= characteristc\_region

FT /note= "see AAW34690"

FT Region 26..29

FT /label= characteristc\_region

FT /note= "see AAW34691"

FT Region 173..177

FT /label= characteristc\_region

FT /note= "see AAW34692"

FT Region 271..274

FT /label= characteristc\_region

FT /note= "see AAW34693"

XX US627061-A.

PN 06-MAY-1997.

XX 07-JUN-1995; 95US-0476008.

XX 13-SEP-1994; 94US-0306063.

PR 31-AUG-1990; 90US-0576537.

PR 28-AUG-1991; 91US-0749611.

XX (MONS ) MONSANTO CO.

XX Barry GF, Kishore GM, Padgett SR, Stallings WC;

PI WPI; 1997-271315/24.

XX N-PSDB; AAT93789.

DR Production of glyphosate-herbicide tolerant plants - using DNA

XX encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase

PT enzyme(s)

XX Disclosure; Fig 4; 151pp; English.

XX AAW34683-89 represent a new class of glyphosate-tolerant

CC 5-enolpyruvylshikimate-3-phosphate synthases (EPSPs). These novel

CC EPSPs enzymes have little homology with known Class I EPSPs enzymes, and

CC belong to a new class, Class II. The present sequence was isolated from

CC *Achromobacter* sp. strain LBAA. The EPSPs enzymes are part of the

CC shikimic acid pathway, which leads to the biosynthesis of aromatic

CC compounds. EPSPs converts phosphoenolpyruvic acid (PEP) and

CC 3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is

CC inhibited by the herbicide glyphosate. It would be useful to produce

CC transgenic crops containing glyphosate resistance genes so that

CC glyphosate-containing herbicides can be applied to selectively kill

CC weeds. The novel EPSPs enzymes exhibit a low Km for PEP and a high Ki for

CC glyphosate-tolerant, and EPSPs enzyme activity is not affected. These

CC class II EPSPs enzymes are fused to a chloroplast transit peptide to

CC target the protein into the chloroplast, which is the site for the

CC shikimic acid pathway. In addition, the EPSPs gene is cloned into a plant

CC under the control of a promoter such as figwort mosaic virus promoter or

CC the cauliflower mosaic virus promoter, so that expression is enhanced.

XX Sequence "449 AA;

SQ

Query Match 83.3%; Score 1906.5; DB 18; Length 449;

Best Local Similarity 83.1%; Pred. No. 1.le-152;

Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASRPATARKSGLSGTVRIPGDKSISHRSMFGGLASGETRITGLLEGEDVINTG 60

DB 1 mshasapxpatarrsealtgeiripgdkslshrsfmfgglasgetritgllegedvintg 60

QY 61 KAMQAMGARIRKSGDTWIIDGVGGLLAPPEAPLDFGNAATGCRLTMGLVGVDFDFTFI 120

DB 61 ramdamgakirkegvdwllngvngccllqpeaaaldfnagtgartlmglvtydmktsf 120

QY 121 GDASLTRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPTTYRVPMASAQVSAV 180

DB 121 gdaslskrpmsgrvlnplremgvqveaagdrmpitllgpktnanpityrvpmasaqvksav 180

QY 181 LLAGINTPGITTVIEPIWTRDHTKMLQGFGANLTVETDADGVTRIRLEGRKLTGOVID 240

DB 181 lllagintpgvttviepvmtrdhtekmlqgfgadltvetcdkgvrhrlrtgggkllvgqclid 240

QY 241 VPGDPSSTAFPLVAALLVPGSDVTTLNVLNMPTRTGLTLTQEMGADIEVINPRLAGGED 300

DB 241 vpgdpsstafplvaallvegsvdvtirnlmnprrtlltqemgadielvinarlagged 300

QY 301 VADLRVRSSTLKGTVVPEDRAPSMIDEYPILAVAAFAEGATVMNGLEELRVKESDRLSA 360

DB 301 vadlrvrasklkgvvvpprapsmideypvllaiaasfaegetvmvgldelrvkesdr 360

QY 361 VANGKLKNGVDCDEGETSLVYVRGPDGKGLGNASCAAAVATHLDHRIAMSFVLMGLVSENP 420

DB 361 vargleangvdcdegtsltvrgprdgkglg---ggvtathldhriamsflvmglaaekp 417

QY 421 VTVDATMIATSFPEFMDLMAGLGAKIELS 450

DB 418 vtvdsmniatsfpefmdmmpglgakiels 447

RESULT 12

AAW34685

ID AAW34685 standard; Protein; 449 AA.

XX AAW34685;

AC AAW34685;

XX 17-FEB-1998 (first entry)

DT Class II EPSP synthase (EPSPs) from *Pseudomonas* sp. strain PG2982.

DE 5-enolpyruvylshikimate-3-phosphate synthase; EPSPs; glyphosate;

XX shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;

XX 5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;

XX glyphosate resistance gene; glyphosate-tolerance; promoter.

XX *Pseudomonas* sp. strain PG2982.

OS Key Location/Qualifiers

XX Region 200..204

FT /label= characteristc\_region

FT /note= "see AAW34690"

FT Region 26..29

FT /label= characteristc\_region

FT /note= "see AAW34691"

FT Region 173..177

FT /label= characteristc\_region

FT /note= "see AAW34692"

FT Region 273..276

FT /label= characteristc\_region

FT /note= "see AAW34693"

XX US627061-A.

PN 06-MAY-1997.

PD







PT useful for characterisation of the enzyme to determine inhibition  
 PT data values

XX Disclosure: Fig 5; 152pp; English.

XX An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the  
 CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),  
 CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,  
 CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,  
 CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,  
 CC produced by recombinant methods, can be used in kinetic studies to  
 CC determine Ki and Km values of the enzyme for its characterisation. The  
 CC enzyme is normally used for the production of 5-enolpyruvyl-3-  
 CC phosphoshikimate acid in plants, and most forms of the enzyme are  
 CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.  
 CC Inhibition data enables more accurate values of concentrations of  
 CC herbicide to be used when growing the plant without being detrimental to  
 CC it. This enables the plant to be grown in the presence of the herbicide,  
 CC being used to inhibit the growth of undesired plants. The present  
 CC sequence represents a Class II EPSPS from bacterial isolate  
 CC Pseudomonas sp. strain PG2982.

XX Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 19; Length 449;  
 Best Local Similarity 83.1%; Pred. No. 1.1e-152;  
 Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

QY 1 MSHGASSRATARKSSGLSGTWRIPGDKISHSRSMFGGLASGETRITGLLEGEDVINTG 60  
 DB 1 mshsaspkpatarsealtgeiripgdkisrsmfmgglasgetritgllegedvintg 60

QY 61 KAMQANGARTKSGDWIWDGVNGGLLAPADLPDFGNATGCRMLGVGVYDFDSTFI 120  
 DB 61 ranqangakirkegdvwiwdgvnggcllqpeaaldfgnagtgarltmglvtydmktsfi 120

QY 121 GDASLTKRPMGRVNLPLREMGVGVKSGEDGRLPVTLRGPKTPITYRVPMASQAQVKSAY 180  
 DB 121 gdaslkrpmpgrvnlplremgvvgeaadgrmpitligpktpantpityrvpmasaqvksav 180

QY 181 LLAGLNTPGITTVIEPIIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRKLTQVID 240  
 DB 181 llaglnpvgttviepvmtrdhtekmlqgfgadltvetdkdgvhrirtggklvgqtd 240

QY 241 VPGDPSSTAPPLVAALLVPGSDVTILNVLNPTRTGLILTLQEMGADIEVNPRLAGGED 300  
 DB 241 vpgdpsstaplvaallvpgsdvtilnvlmnptrtgliltlqemgadievlnarlagged 300

QY 301 VADLRVRSSTLKGVTVPEDRAPSMIDEPYILAVAAFAEGATVMNGLEELRVKESDRLSA 360  
 DB 301 vadlrvrssstlkgvtpedrapsmidpypilavaaafegatvmngleelrvkesdrlsa 360

QY 361 VANGKLNGVDCDEGETSLVVRGPRPGKGLGNASGAAVATHLDHRIAMSLVMGLVSENP 420  
 DB 361 vangleangvdcdegetslvvrgrprpgkglgnasgaavathldhriamslvmglvse 420

QY 421 VTVDATMIATSPFPMMDLMLAGLAKIELS 450  
 DB 418 vtvdsmiatstfpmfmdmmpglgakiels 447

Search completed: August 15, 2002, 13:58:47  
 Job time: 66 sec

**THIS PAGE BLANK (USPTO)**